AF IZW

Attorney's Docket No.: 13425-047001 / 00357-US

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant	: Johan Weigelt et al.	Art Unit	: 1641
Serial No.	: 09/986,240	Examiner	: Deborah A. Davis
Filed	: October 19, 2001		
Title	: NEW NUCLEAR MAGNETIC RESONANCE SCREENING METHOD		

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

APPEAL BRIEF

Appellants submit this Appeal Brief to the Board of Patent Appeals and Interferences in response to the final Office Action dated May 4, 2004. A Notice of Appeal was filed on October 4, 2004 and received in the Patent and Trademark Office on October 7, 2004, making this Appeal Brief due on December 7, 2004.

The $340 fee for filing the Appeal Brief fee is enclosed. Please apply any other charges or credits to Deposit Account No. 06-1050, referencing Attorney Docket No. 13425-047001.

(I) Real Party in Interest

The real party in interest is the assignee, Biovitrum AB.

(II) Related Appeals and Interferences

There are no appeals or interferences for related cases.

CERTIFICATE OF MAILING BY FIRST CLASS MAIL

I hereby certify under 37 CFR §1.8(a) that this correspondence is being deposited with the United States Postal Service as first class mail with sufficient postage on the date indicated below and is addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

DECEMBER 7, 2004
Date of Deposit

Carol Tyrrell
Signature

CAROL TYRRELL
Typed or Printed Name of Person Signing Certificate

12/13/2004 CCHAU1 00000005 09986240
01 FC:1402 340.00 OP

(III) Status of Claims

Claims 1-10 and 12 are pending and were all finally rejected in an Office Action dated May 4, 2004 ("the Action"). Claims 1-10 and 12 are the subject of the present appeal. A copy of the appealed claims is attached to this brief in a Claims Appendix.

(IV) Status of Amendments

No amendments have been filed subsequent to the final rejection.

(V) Summary of Claimed Subject Matter

Independent claim 1 is directed to a method for identifying at least one binder molecule that interacts with a polypeptide or protein. The claimed method recites six steps, (a) through (f) (specification at page 2, line 26, to page 3, line 16). In step (a), two amino acid types (AA1 and AA2) in a polypeptide or protein of interest are chosen, whereby AA2 at least once occurs directly subsequent to AA1 in the amino acid sequence of the polypeptide or protein (specification at page 2, line 28, to page 3, line 2). In step (b), the two amino acid types in the polypeptide or protein are labeled with ^{13}C (AA1) and ^{15}N (AA2) (specification at page 3, lines 3-5). In step (c), a first NMR spectrum of the labeled polypeptide or protein is generated, which spectrum identifies signals from the labeled amino acid pair (specification at page 3, lines 6-8). In step (d), the labeled polypeptide or protein is contacted with a potential binder molecule or a mixture of potential binder molecules having a molecular mass of from 50 to 1000 Da (specification at page 6, lines 19-21). Contacting step (d) is carried out under conditions and for sufficient time so as to allow binding between the polypeptide or protein and the potential binder molecule(s) (specification at page 3, lines 9-11). In step (e), an NMR spectrum from the mixture of step (d) is generated (specification at page 3, lines 12-13). In step (f), the first and second NMR spectra are compared to identify a chemical shift change between the two spectra, wherein a shift indicates an interaction between the potential binder molecule and the labeled polypeptide or protein (specification at page 3, lines 14-16).

Dependent claim 12 further limits claim 1 by requiring that the method be used for screening a compound library (specification at page 4, lines 12-13).

(VI) Grounds of Rejection to be Reviewed on Appeal

(A) Claims 1, 4, 5, 8-10, and 12 stand rejected under 35 U.S.C. § 103(a) as allegedly unpatentable over Yabuki et al. (1998) J. BIOMOLECULAR NMR, 11:295-306 (hereinafter "Yabuki") in view of Moore, U.S. Patent No. 6,060,603.

(B) Claims 2, 3, 6, and 7 stand rejected under 35 U.S.C. § 103(a) as allegedly unpatentable over Yabuki in view of Moore and in further view of Fesik et al., WO 97/18471 (hereinafter "Fesik").

(VII) Argument

(A) Rejection of claims 1, 4, 5, 8-10, and 12 under 35 U.S.C. § 103(a) as allegedly unpatentable over Yabuki in view of Moore

To establish a *prima facie* case of obviousness under 35 U.S.C. § 103(a), there must be some suggestion or motivation, either in the cited references or in the knowledge generally available to one of ordinary skill in the art, to modify a single reference or to combine it with one or more other references. The suggestion or motivation to make the claimed combination must be found in the prior art and not based on applicant's disclosure. *In re Vaeck,* 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991); MPEP § 706.02(j).

Yabuki describes a dual amino acid-selective labeling method and its use for investigating structures of proteins and protein complexes. Yabuki does not suggest using the labeling technique to carry out a screen to identify whether a candidate molecule has the ability to bind to a given target protein. Instead, Yabuki describes the use of the labeling technique to investigate the nature of the interaction between two proteins (Ras and Raf) that were previously known to bind to each other.

The Action concedes that Yabuki does not disclose carrying out its labeling method with a binder molecule having a molecular mass of from 50 to 1000 Da (the Action at page 3).

Moore describes methods for modeling biologically active ligands and designing mimetics of such ligands. The Action alleges that it would have been obvious to one of ordinary skill in the art to modify the disclosure of Yabuki to include the use of small molecular weight ligands taught by Moore (the Action at page 4).

The person of ordinary skill in the art, at the time the present application was filed, would have lacked the requisite suggestion or motivation to modify Yabuki in the manner contemplated by the Action. Moore does not describe any molecule having a molecular mass of from 50 to 1000 Da that has the ability to bind to Ras or any other protein described in Yabuki. Furthermore, nothing in Moore suggests modifying the protein structure analysis method of Yabuki to instead carry out a screening method that compares a first NMR spectra (of a labeled polypeptide or protein) and a second NMR spectra (of the labeled polypeptide or protein that has been contacted with a potential binder molecule having a molecular mass of from 50 to 1000 Da).

Because Yabuki is concerned with investigations of protein structure and the characterization of protein complexes, the skilled artisan would have had no reason to modify the methods of Yabuki to create a method that screens candidate compounds to evaluate their ability to bind to Ras or any other protein. Furthermore, neither Yabuki nor Moore describe a single low molecular weight compound (i.e., having a molecular mass of from 50 to 1000 Da, as is required by the claims) that was known to bind to Ras that the skilled artisan even might have attempted to use in the method of Yabuki in place of Raf to investigate the nature of a Ras-ligand interaction.

The Action stated that "absent evidence to the contrary, the range recited in the instant claims from 50-1000 Daltons is viewed as mere optimization of the prior art assay" (the Action at page 4). Because Yabuki does not describe a screening assay to identify potential binder molecules (but instead describes experiments that evaluate Ras binding to Raf), there would have been no basis or rationale for the skilled artisan to "optimize" the methods of Yabuki by, for example, replacing Raf (a known ligand of Ras) with some other molecule having a molecular mass of from 50 to 1000 Da. Yabuki describes a dual amino acid-selective labeling method using a protein (Ras) and a known binding partner (Raf) and thus gives no suggestion or motivation to replace Raf with low molecular weight compounds having no known Ras-binding ability.

Accordingly, Appellants respectfully submit that the Action does not establish a *prima facie* showing of obviousness and request that the rejection of independent claim 1 and dependent claims 4, 5, 8-10, and 12 be withdrawn.

(B) Rejection of claims 2, 3, 6, and 7 under 35 U.S.C. § 103(a) as allegedly unpatentable over Yabuki in view of Moore and in further view of Fesik

The Action contends that it would have been obvious to one of ordinary skill in the art to incorporate a comparison method of the various assays as taught by Fesik into the method of Yabuki in view of Moore to compare the binding of ligands to various biomolecules determined by NMR and to also observe chemical shifts detected by 2-D NMR techniques (the Action at page 5).

As detailed in Section (A) above, Yabuki describes the use of a dual amino acid-selective labeling method to investigate structures of proteins and protein complexes. Nothing in Yabuki suggests generating the two NMR spectra recited in the claims and subsequently comparing the two spectra to identify an interaction between a potential binder molecule having a molecular mass of from 50 to 1000 Da and a labeled polypeptide. As detailed above, Moore provides no suggestion or motivation to modify the technique of Yabuki to convert the method described therein into an assay to screen for interactions between a labeled polypeptide and a potential binder molecule having a molecular mass of from 50 to 1000 Da. Accordingly, the combination of Yabuki and Moore do not render obvious claim 1 or the claims that depend therefrom.

Fesik does not add what is lacking in Yabuki and Moore. Fesik's disclosure of processes for identifying compounds would not have provided the requisite suggestion or motivation to cause the skilled artisan to so thoroughly alter the methods described by Yabuki so as to, instead of characterizing protein structures as taught in Yabuki, modify the labeling techniques described therein to be used for the altogether different purpose of carrying out a screening assay to identify potential binder molecules. Yabuki is not directed to screening assays and nothing in Moore or Fesik, considered either alone or in combination, would have led the skilled artisan to (a) carry out the methods of Yabuki with a molecule having a molecular mass of from 50 to 1000 Da, and/or (b) fundamentally modify the techniques of Yabuki so as to screen to identify binder molecules.

In light of these comments, Appellants respectfully submit that the Action does not establish a *prima facie* showing of obviousness and request that the rejection of dependent claims 2, 3, 6, and 7 be withdrawn.

(C) Rejection of claim 12 under 35 U.S.C. § 103(a) as allegedly unpatentable over Yabuki in view of Moore

Claim 12 was not rejected under a separate heading in the Action. However, appellants consider claim 12 to be patentable for the following separate reasons, in addition to those reasons provided in Section (A) above.

Dependent claim 12 requires that the method of independent claim 1 be used for screening a "compound library." Nothing in Yabuki or Moore, taken alone or in combination, suggests applying the dual amino acid-selective labeling method of Yabuki to a "compound library" to identify molecules in the library that bind to a polypeptide or protein of interest. In particular, there is no description whatsoever in Moore of any compound library that the skilled artisan would have had reason to use in a labeling method of Yabuki.

Accordingly, Appellants respectfully submit that the Action does not establish a *prima facie* showing of obviousness and request that the rejection of dependent claim 12 be withdrawn.

For the foregoing reasons, Appellants respectfully submit that the Action's rejections of claims 1-10 and 12 are unwarranted. Appellants therefore request that the Board overturn the rejections.

Respectfully submitted,

Date: December 7, 2004

Jack Brennan
Reg. No. 47,443

Fish & Richardson P.C.
Citigroup Center
52nd Floor
153 East 53rd Street
New York, New York 10022-4611
Telephone: (212) 765-5070
Facsimile: (212) 258-2291

30211243.doc

Claims Appendix

1. A method for identifying at least one binder molecule comprising the steps of:

(a) choosing two amino acid types (AA1 and AA2) in a polypeptide or protein of interest, whereby AA2 at least once occurs directly subsequent to AA1 in the amino acid sequence of the polypeptide or protein, defining an amino acid pair AA1-AA2;

(b) labeling the two amino acid types (AA1 and AA2) in the polypeptide or protein of interest, whereby all AA1-residues is labeled with ^{13}C and all AA2-residues with ^{15}N;

(c) generating a first HNCO-type NMR spectrum of the labeled polypeptide or protein from step (b), thereby identifying signals from the labeled amino acid pair AA1-AA2;

(d) contacting the labeled polypeptide or protein with a potential binder molecule or a mixture of potential binder molecules under conditions and sufficient time for allowing binding of the potential binder molecule or the potential binder molecules and the labeled polypeptide or protein, wherein the potential binder molecule or potential binder molecules have a molecular mass of from 50 to 1000 Da;

(e) generating a second HNCO-type NMR spectrum, or a ^{1}H-^{15}N correlation type NMR spectrum, of the mix from step (d), monitoring signals identified in step (c); and

(f) comparing the first and the second NMR spectra, whereby a chemical shift change of the signals identified in step (c) between the two spectra indicates an interaction between the potential binder molecule or the potential binder molecules and the labeled polypeptide or protein.

2. The method of claim 1, wherein the labeled amino acid pair AA1-AA2 is unique within a sphere radius of 10 Å within the polypeptide or protein, wherein the 10 Å radius is measured from the labeled amino acid pair AA1-AA2.

3. The method of claim 1, wherein the labeled amino acid pair AA1-AA2 is unique within a sphere radius of 50 Å within the polypeptide or protein, wherein the 50 Å radius is measured from the labeled amino acid pair AA1-AA2.

4. The method of claim 1, wherein the labeled amino acid pair AA1-AA2 is unique within the polypeptide or protein.

5. The method of claim 1, wherein the labeled amino acid pair AA1-AA2 is within a binding pocket of the polypeptide or protein.

6. The method of claim 1, wherein the labeled amino acid pair AA1-AA2 is in the proximity of an active site within the polypeptide or protein.

7. The method of claim 1, wherein the result of the method is compared to the result of any other suitable binding or activity assay.

8. The method of claim 1, wherein the polypeptide or protein has a size of from 10 to 150 kDa.

9. The method of claim 1, wherein the potential binder molecule is a peptide, polypeptide, protein, antibody, nucleic acid molecule, or carbohydrate.

10. The method of claim 1, wherein the potential binder molecule comprises a complex of one or more of a peptide, polypeptide, protein, antibody, nucleic acid molecule, or carbohydrate.

12. The method of claim 1, wherein the method is used for screening a compound library.

Evidence Appendix

Enclosed are copies of the following evidence relied upon by the Examiner as to grounds of rejection to be reviewed on appeal.

Exhibit 1: Yabuki et al. (1998) J. BIOMOLECULAR NMR, 11:295-306

Exhibit 2: Moore, U.S. Patent No. 6,060,603

Exhibit 3: Fesik et al., WO 97/18471

Related Proceedings Appendix

There are no appeals or interferences for related cases.

Journal of Biomolecular NMR, 11: 295–306, 1998.
KLUWER/ESCOM

Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis

Takashi Yabuki[a,b,*], Takanori Kigawa[a,*], Naoshi Dohmae[c], Koji Takio[c], Tohru Terada[a,b,c], Yutaka Ito[d], Ernest D. Laue[e], Jonathan A. Cooper[f], Masatsune Kainosho[g] & Shigeyuki Yokoyama[a,b,**]

[a]*Cellular Signaling Laboratory, the Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01, Japan;* [b]*Department of Biophysics and Biochemistry, Graduate School of Science, the University of Tokyo, Bunkyo-ku, Tokyo 113, Japan;* [c]*Division of Biomolecular Characterization and* [d]*Laboratory of Cellular and Molecular Biology, the Institute of Physical and Chemical Research (RIKEN), Wako, Saitama 351-01, Japan;* [e]*Cambridge Centre for Molecular Recognition, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, U.K.;* [f]*Fred Hutchinson Cancer Research Center, Seattle, WA 98104, U.S.A.;* [g]*Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, Hachioji, Tokyo 192-03, Japan*

Received 7 August 1997; Accepted 20 October 1997

Key words: cell-free protein synthesis, selective labeling, site-directed labeling, stable isotope

Abstract

We developed two methods for stable-isotope labeling of proteins by cell-free synthesis. Firstly, we applied cell-free synthesis to the dual amino acid-selective ^{13}C–^{15}N labeling method, originally developed for *in vivo* systems by Kainosho and co-workers. For this purpose, we took one of the advantages of a cell-free protein synthesis system; the amino acid-selective stable-isotope labeling is free of the isotope scrambling problem. The targets of selective observation were Thr^{35} and Ser^{39} in the 'effector region' (residues 32–40) of the Ras protein complexed with the Ras-binding domain of c-Raf-1 (Raf RBD) (the total molecular mass is about 30 kDa). Using a 15-mL *Escherichia coli* cell-free system, which was optimized to produce about 0.4 mg of Ras protein per 1-mL reaction, with 2 mg each of DL-[$^{13}C'$]proline and L-[^{15}N]threonine, we obtained about 6 mg of Ras protein. As the Pro–Thr sequence is unique in the Ras protein, the Thr^{35} cross peak of the Ras•Raf RBD complex was unambiguously identified by the 2D ^{1}H–^{15}N HNCO experiment. The Ser^{39} cross peak was similarly identified with the [$^{13}C'$]Asp/[^{15}N]Ser-selectively labeled Ras protein. There were no isotope scrambling problems in this study. Secondly, we have established a method for producing a milligram quantity of site-specifically stable-isotope labeled protein by a cell-free system involving amber suppression. The *E. coli* amber suppressor $tRNA^{Tyr}_{CUA}$ (25 mg) was prepared by *in vitro* transcription with T7 RNA polymerase. We aminoacylated the $tRNA^{Tyr}_{CUA}$ transcript with purified *E. coli* tyrosyl-tRNA synthetase, using 2 mg of L-[^{15}N]tyrosine. In the gene encoding the Ras protein, the codon for Tyr^{32} was changed to an amber codon (TAG). This template DNA and the [^{15}N]Tyr-$tRNA^{Tyr}_{CUA}$ were reacted for 30 min in 30 mL of *E. coli* cell-free system. The subsequent purification yielded 2.2 mg of [^{15}N]Tyr^{32}-Ras protein. In the ^{1}H–^{15}N HSQC spectrum of the labeled Ras protein, only one cross peak was observed, which was unambiguously assigned to Tyr^{32}.

*The first two authors contributed equally to this work.

**To whom correspondence should be addressed. Mailing address: Cellular Signaling Laboratory, the Institute of Physical and Chemical Research (RIKEN), 2-1 Hirosawa, Wako, Saitama 351-01, Japan.

Abbreviations: CK, creatine kinase; CP, creatine phosphate; folinic acid, L(−)-5-formyl-5,6,7,8-tetrahydrofolic acid; GMPPNP, guanosine 5′-*O*-(β,γ-imidotriphosphate); MALDI-TOF, matrix assisted laser desorption ionization-time of flight; PEG, polyethylene glycol; Raf RBD, Ras-binding domain of c-Raf-1; Rap1A•GMPPNP, GMPPNP-bound Rap1A; Ras•GDP, GDP-bound Ras protein; Ras•GMPPNP, GMPPNP-bound Ras protein; SDIL, site-directed isotope labeling; $TyrRS_t$, tyrosyl-tRNA synthetase.

Introduction

Protein biosynthesis systems, such as the *Escherichia coli* coupled transcription-translation system, have been remarkably improved, and a milligram quantity of protein can be prepared (Kigawa et al., 1995; Kim et al., 1996). For NMR spectroscopy, cell-free protein synthesis has a number of advantages over the ordinary recombinant protein synthesis *in vivo*. We have been applying the cell-free protein synthesis to stable-isotope labeling (Kigawa et al., 1995).

Kainosho and co-workers have developed a dual amino acid-selective ^{13}C–^{15}N labeling technique (Kainosho and Tsuji, 1982; Kainosho et al., 1985, 1987; Westler et al., 1988a,b). This technique utilizes protein samples in which the main chain carbonyl carbons of one amino acid type are labeled with ^{13}C and the amide nitrogens of another amino acid type are labeled with ^{15}N. The NMR signals of the amino acid residues that possess a ^{13}CO–^{15}N linkage can be extracted on the basis of the ^{13}C–^{15}N spin coupling. This technique was applied to an investigation of the structure of *Streptomyces* subtilisin inhibitor and its complex with the target enzyme, subtilisin (Kainosho and Tsuji, 1982; Kainosho et al., 1985, 1987; Westler et al., 1988a,b), and was also applied to structural analyses of an antibody and its fragment (for a review, see Arata et al., 1994). As indicated in these studies, the advantage of this dual labeling technique is that sequence specific assignments can be performed, even for very large proteins, such as an entire IgG molecule (molecular mass about 150 kDa, Kato et al., 1989). However, if *in vivo* expression methods are used for this kind of amino acid-selective labeling, then the amino acid metabolism can cause an isotope-scrambling problem in some cases, such as for Asp or Ser residues (Kainosho and Tsuji, 1982; Kainosho et al., 1987; McIntosh and Dahlquist, 1990). This is a serious drawback of this technique.

This scrambling problem can be solved by using a cell-free protein synthesis system. With this system, the Asp residues could be labeled without labeling the Asn residues, and similarly, the Ser residues could be labeled without labeling the Gly residues (Kigawa et al., 1995), whereas *in vivo* expression methods do not allow this selectivity (McIntosh and Dahlquist, 1990; Yamasaki et al., 1992). In this study, we used the cell-free system for dual labeling to investigate the structure of the effector region of the human c-Ha-Ras protein complexed with the Ras-binding domain of the rat c-Raf-1 protein (Raf RBD; the total molecular mass of the complex is about 30 kDa). This allowed some of the main-chain amide signals of the effector region to be assigned in a sequence specific manner by the HNCO experiments.

In the next step, we tried site-directed incorporation of a stable-isotope labeled amino acid into a protein, which would be useful for NMR analyses. The site-directed incorporation of unnatural amino acids was achieved by cell-free protein synthesis involving amber suppression (Noren et al., 1989; Bain et al., 1991). Ellman et al. (1992) produced T4 lysozyme containing a unique ^{13}C-labeled Ala using chemically aminoacylated amber suppressor tRNA, and recorded ^{13}C-filtered proton NMR spectra. The yield was 0.15 mg of lysozyme from a 40-mL reaction mixture.

Sonar et al. (1994) and Liu et al. (1995) reported the site-directed labeling of bacteriorhodopsin with ^{2}H and ^{13}C, respectively; an overproduced amber suppressor, $tRNA^{Tyr}_{CUA}$, from *E. coli* was aminoacylated with a stable-isotope labeled L-tyrosine by an *E. coli* crude protein fraction, and was used in the cell-free translation of a messenger RNA with an amber codon (UAG) at a specified site. The yield was 0.06 mg of bacteriorhodopsin from a 2-mL reaction mixture.

In the present study, an *E. coli* cell-free system was applied to the site-directed stable-isotope labeling of the Ras protein. An *E. coli* suppressor $tRNA^{Tyr}_{CUA}$ was prepared by *in vitro* transcription with T7 RNA polymerase, and was aminoacylated with purified tyrosyl-tRNA synthetase. The aminoacylated $tRNA^{Tyr}_{CUA}$ was used in the cell-free coupled transcription-translation of a template DNA with an amber codon (TAG) for Tyr^{32}. We obtained 2.2 mg of purified [^{15}N]Tyr^{32}-Ras protein and succeeded in measuring the ^{1}H–^{15}N HSQC spectrum.

Materials and methods

Preparation of the selectively $^{13}C'/^{15}N$-labeled Ras protein

The Ras protein used in this study consisted of 171 amino acid residues, and lacked the C-terminal 18 amino acid residues (Ha et al., 1989). The *E. coli* S30 extract for the cell-free protein synthesis was prepared according to Pratt (1984) from *E. coli* strain A19 (*metB*, *rna*). The T7 RNA polymerase was prepared according to Zawadzki and Gross (1991). The large scale reaction for selective $^{13}C'/^{15}N$-labeling of the Ras protein was performed by our recently improved

protocol (to be published elsewhere). The reaction mixture (15 mL final volume) consisted of 1.0 mM each of DL-[^{13}C']proline (M. Kainosho, unpublished) and L-[^{15}N]threonine (Isotec), 1.0 mM each of the other 18 amino acids, 6.7 μg/mL of the pK7-Ras plasmid with the T7 promoter and the structural gene for the Ras protein (Kigawa et al., 1995), 55 mM Hepes-KOH (pH 7.5), 1.7 mM DTT, 1.2 mM ATP, 0.8 mM each of CTP, GTP, and UTP, 80 mM creatine phosphate (CP) (Boehringer-Mannheim), 250 μg/mL creatine kinase (CK) (Boehringer-Mannheim), 4.0% polyethylene glycol (PEG) 8000 (Sigma), 0.64 mM 3',5'-cyclic AMP, 68 μM L(−)-5-formyl-5,6,7,8-tetrahydrofolic acid (folinic acid), 175 μg/mL *E. coli* total tRNA (Boehringer-Mannheim), 210 mM potassium glutamate, 27.5 mM ammonium acetate, and 10.7 mM magnesium acetate. These components were mixed and preincubated at 37 °C for 5 min, after which the S30 extract (300 μL/mL) and T7 RNA polymerase (93 μg/mL) were added. The reaction was incubated at 37 °C for 1 h, and was then centrifuged for 5 min at 12000 × g. The supernatant was desalted on a PD-10 gel-filtration column (Pharmacia). The Ras protein was purified using an FPLC Mono Q anion exchange column (Pharmacia) followed by an FPLC Superdex 75 gel-filtration column (Pharmacia). The yield of the Ras protein was determined from the specific absorbance: A_{280} = 0.56 cm^{-1} mg^{-1} mL (Yamasaki et al., 1992). The large scale reaction for [^{13}C']Asp/[^{15}N]Ser-selective labeling was performed in the same manner, except L-[^{13}C']aspartate (M. Kainosho, unpublished) and L-[^{15}N]serine (Isotec) were used in place of DL-[^{13}C']proline and L-[^{15}N]threonine.

Preparation of the Raf RBD

For the Raf RBD preparation, we used the *E. coli* strain BL21(DE3), which produces the rat Raf-1 RBD (amino acids 51–131) containing the extra vector derived residues MASMTGGQQMGRGS and KLAAALEHHHHHH as its N- and C-terminal residues, respectively. The Raf RBD was purified from the lysed cell extract by Ni-NTA agarose Ni-chelate affinity chromatography (Qiagen) followed by CM TOYOPEARL cation exchange chromatography (Tosoh). The yield of the Raf RBD was determined by measuring the tryptophan absorbance at 280 nm.

The template DNA for in vitro preparation of the suppressor tRNA$^{Tyr}_{CUA}$

We constructed a fragment, with the T7 promoter, the gene for the amber suppressor tRNA$^{Tyr}_{CUA}$, and a *Bst*NI cleavage site, from the cloned *E. coli* wild-type tRNA$^{Tyr}_2$ gene by polymerase chain reaction with three synthetic primers. The fragment was ligated into pUC119 (pTRY1), and its nucleotide sequence was confirmed with an ALFred DNA sequencer (Pharmacia). The pTRY1 plasmid DNA (1 mg) was digested in a reaction mixture (5 mL) with *Bst*NI (1500 units) at 55 °C for 17 h, and was phenol-chloroform extracted, ethanol-precipitated, and dried.

In vitro *preparation of the suppressor tRNA$^{Tyr}_{CUA}$*

The transcription was carried out in a reaction mixture (10 mL) containing 40 mm Hepes-KOH (pH 8.1), 46 mM $MgCl_2$, 5 mM DTT, 40 μg/mL BSA, 2 mM spermine, 0.01% Triton-X100, 6.25 mM each of ATP, CTP, GTP, and UTP, 20 mM 5'-GMP, 0.1 mg/mL BstNI-digested template DNA, 20 units/mL of RNase inhibitor (Takara), and 60 μg/mL T7 RNA polymerase. After an incubation at 40 °C for 5 h, 1 mL of 500 mM EDTA (pH 8.0) was added, and the reaction mixture was desalted on a PD-10 column. The transcribed tRNA was then purified by DEAE-Toyopearl column chromatography (Tosoh) followed by electrophoresis on a 20% polyacrylamide gel containing 7M urea.

Preparation of the E. coli tyrosyl-tRNA synthetase (TyrRS)

The *E. coli tyrS* gene was cloned from the *E. coli* strain HB101 chromosome by polymerase chain reaction, and was ligated into the T7 promoter expression vector pT7-7 (pT7-7-tyrS). The *E. coli* strain JM109 (DE3) (Promega) harboring pT7-7-tyrS was cultured at 37 °C in M9ZB medium (Studier et al., 1990) containing 50 μg/mL ampicillin. The TyrRS protein was purified from the lysed cell extract by DEAE-Sephacel column chromatography (Pharmacia) followed by FPLC Phenyl Superose column chromatography (Pharmacia). The yield of the purified TyrRS protein was determined by a protein assay (Bio-Rad).

Small-scale preparation of the Tyr-tRNA$^{Tyr}_{CUA}$

The aminoacylation of the *E. coli* suppressor tRNA$^{Tyr}_{CUA}$, prepared as described above, was carried out in a reaction mixture (60 μL) containing 50 mM Hepes-KOH (pH 7.8), 10 mM $MgCl_2$, 4 mM ATP, 133 μM tRNA$^{Tyr}_{CUA}$, 280 μM L-[^{14}C]tyrosine

(2.23 GBq/mmol, Moravek), and 2 μM of the purified TyrRS. After an incubation at 37 °C for 30 min, the reaction was stopped by the addition of 1.2 μL of a solution (pH 4.5) containing 12.4 M acetic acid and 3 M ammonium acetate. The [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$, together with the uncharged tRNA$^{Tyr}_{CUA}$, was phenol-chloroform extracted, and then desalted on a BioSpin-6 gel-filtration column (Bio-Rad). The sample was divided into seven tubes, ethanol-precipitated, and dried. Two tubes were immediately used for the two types of small-scale cell-free protein synthesis experiments (with unlabeled or ^{14}C-labeled L-leucine, as described below). Another tube was used to determine the amount of [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ by liquid scintillation counting of the trichloroacetic acid insoluble material.

The template DNA for site-directed incorporation

The pK7-Ras plasmid was mutated at the Tyr^{32} codon (TAC) of the Ras protein to an amber codon (TAG) (pK7-RasY32am). The nucleotide sequence was confirmed with an ALFred DNA sequencer.

Small-scale cell-free reaction for site-directed incorporation

The pellet of the [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ sample in a tube (described above) was dissolved to a final concentration of 40 μM in a solution (30 μL) containing 55 mM Hepes-KOH (pH 7.5), 2 mM DTT, 1.2 mM ATP, 0.8 mM each of CTP, GTP, and UTP, 80 mM CP, 250 μg/mL CK, 4.0% PEG 8000, 0.64 mM 3′,5′-cyclic AMP, 68 μM folinic acid, 175 μg/mL *E. coli* total tRNA, 210 mM potassium glutamate, 27.5 mM ammonium acetate, 10.7 mM magnesium acetate, 0.64 mM unlabeled L-leucine, 1 mM L-tyrosine, 0.7 mM each of the other 18 amino acids, 6.7 μg/mL of the pK7-RasY32am plasmid, 133 μg/mL T7 RNA polymerase, and 7.2 μL S30 extract.

This reaction mixture was incubated at 37 °C for 1 h. At 5, 10, 20, 30, 45, and 60 min, 3.7 μL aliquots of the reaction mixture, containing both the synthesized protein labeled with L-[^{14}C]tyrosine and the remaining L-[^{14}C]Tyr-tRNA, were removed, mixed with 60 μL of 0.1N NaOH, and kept at 25 °C for 30 min to deacylate the radioactive aminoacyl-tRNA. The incorporation of L-[^{14}C]tyrosine into the Ras protein was determined by liquid scintillation counting of the trichloroacetic acid insoluble material. The amount of the [^{14}C]Tyr-containing Ras protein was estimated based on the assumption that the L-[^{14}C]tyrosine was incorporated only into the position of Tyr^{32}. The amount of residual [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ was estimated from the difference between the L-[^{14}C]tyrosine incorporations with and without the deacylation treatment.

The incorporation of L-[^{14}C]leucine into the Ras protein was determined by subtraction of the L-[^{14}C]tyrosine incorporation from the total incorporation of L-[^{14}C]tyrosine and L-[^{14}C]leucine by the same reaction, except for the use of 0.64 mM L-[^{14}C]leucine (193 MBq/mmol, Amersham) in place of the unlabeled L-leucine in the reaction mixture. The amount of the [^{14}C]Leu-containing Ras protein was estimated on the assumption that the L-[^{14}C]leucine was incorporated into all Leu residues.

In addition, the Ras protein was synthesized by the cell-free coupled transcription-translation system, in the same reaction mixture as that of the site-directed L-[^{14}C]tyrosine incorporation, except for the use of the pK7-Ras plasmid, which lacks the amber mutation, in place of pK7-RasY32am, 0.64 mM L-[^{14}C]leucine (193 MBq/mmol) in place of the unlabeled L-leucine, and no [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$.

These cell-free reaction products were also analyzed by SDS-PAGE with tricine buffer (Schagger and von Jagow, 1987), followed by autoradiography with a Bio Image Analyzer BAS-2000 system (Fuji Film).

Large-scale preparation of [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$

For the synthesis of the [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$, 2.5 mL of 50 mM Hepes-KOH buffer (pH 7.8), containing 10 mM $MgCl_2$, 6 mM ATP, 430 μM tRNA$^{Tyr}_{CUA}$, and 0.6 μM TyrRS, was prepared and saturated with L-[^{15}N]tyrosine (Icon) (2 mg of L-[^{15}N]tyrosine powder was added to the buffer, but was not completely dissolved). The mixture was incubated at 37 °C for 30 min, and the reaction was stopped by the addition of 50 μL of a solution (pH 4.5) containing 12.4 M acetic acid and 3 M ammonium acetate. The mixture was phenol-chloroform extracted, desalted on a PD-10 column, ethanol-precipitated, and dried. The amount of synthesized Tyr-tRNA$^{Tyr}_{CUA}$ was determined by the same aminoacylation reaction in a 50-μL volume, except for the use of 0.92 mM L-[^{14}C]tyrosine (1.78 GBq/mmol) in place of L-[^{15}N]tyrosine, followed by liquid scintillation counting of the trichloroacetic acid insoluble material.

Preparation of [^{15}N]Tyr^{32}-Ras protein using the cell-free protein synthesis system

The [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$ pellet, prepared as described above, was dissolved to a final concentration of 36 μM in a solution (30 mL) containing 55 mM Hepes-KOH (pH 7.5), 2 mM DTT, 1.2 mM ATP,

0.8 mM each of CTP, GTP, and UTP, 80 mM CP, 250 μg/mL CK, 4.0% PEG 8000, 0.64 mM 3′,5′-cyclic AMP, 68 μM folinic acid, 175 μg/mL *E. coli* total tRNA, 210 mM potassium glutamate, 27.5 mM ammonium acetate, 10.7 mM magnesium acetate, 10.8 mg L-tyrosine, 0.7 mM each of the other 19 amino acids, 6.7 μg/mL of the pK7-RasY32am plasmid, 133 μg/mL T7 RNA polymerase, and 7.2 mL S30 extract. This reaction mixture was incubated at 37 °C for 30 min, and was desalted on a PD-10 column. The labeled Ras protein was purified to homogeneity with Resource Q (Pharmacia) and Superdex 75 FPLC columns. Samples from each purification step were analyzed by SDS-PAGE. The yield of the [^{15}N]Tyr^{32}-Ras protein was estimated to be 2.2 mg by the protein assay. The unlabeled Ras protein was prepared as described previously (Ha et al., 1989).

Mass spectrometry

The purified Ras proteins were desalted by ultrafiltration using a Centricon-10 unit (Amicon; *Mr* cutoff = 1×10^4) as described (Akashi et al., 1996). The matrix assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectra were then measured and analyzed on a Bruker REFLEX mass spectrometer with a delayed extraction system. The Ras protein samples were analyzed in the linear mode with 3,5-dimethoxy-4-hydroxy-cinnamic acid as the matrix.

In order to determine the level of ^{15}N enrichment of the [^{13}C′]Asp/[^{15}N]Ser selectively labeled Ras protein, the protein was digested with trypsin. One of the digested fragments consisted of residues 136–148, containing two Ser residues (Ser^{136} and Ser^{145}) and no Asp residue. This fragment was analyzed in the reflector mode with 2,5-dihydroxy-benzoic acid as the matrix. The simulation of the isotopic distribution pattern for Ras(136–148) was carried out with the Bruker Xmass program on a SUN SS5 workstation. The patterns for the various ^{15}N contents were calculated as linear combinations of the simulated patterns for the unlabeled Ras(136–148) and the labeled Ras(136–148) with two [^{15}N]Ser residues. To estimate the ^{15}N content of Tyr^{32} of the [^{15}N]Tyr^{32}-Ras protein, the protein was digested with *Achromobacter lyticus* protease I (lysylendopeptidase). The digest containing the Ras fragment consisting of residues 17–42, Ras(17–42), was analyzed as described above. The simulation of the isotopic distribution pattern for Ras(17–42) was carried out as described above.

NMR samples

For NMR measurements of selectively ^{13}C′/^{15}N-labeled Ras proteins, 6 mg of each of the proteins, prepared as described above, was used. To prepare the guanosine 5′-*O*-(β,γ-imidotriphosphate)-bound form of the Ras protein (Ras•GMPPNP; GMPPNP = a slowly hydrolyzable analog of GTP), the bound GDP was exchanged by incubating the Ras protein (0.5 mM) three times with GMPPNP (Boehringer-Mannheim, final concentration 2 mM) in the presence of 10 mM EDTA and 10 units/mL apyrase (Sigma) for 10 min at 37 °C. The protein samples were then concentrated and mixed with 6 mg of Raf RBD (Ras•GMPPNP: Raf RBD ≈ 1:2) in NMR buffer (90% H_2O and 10% D_2O containing 20 mM sodium phosphate buffer (pH 6.5), 5 mM DTT, 5 mM $MgCl_2$, 100 mM NaCl, and 0.01% NaN_3) using a Centricon-3 unit (Amicon; *Mr* cutoff = 3×10^3).

For an NMR measurement of [^{15}N]Tyr^{32}-Ras protein, 2.2 mg of the protein was used. The solution of the Ras protein (0.2 mL) was mixed with 5 mL of the NMR buffer described above. The solution was concentrated to about 0.2 mL by ultrafiltration using a Centricon-10 unit (Amicon; *Mr* cutoff = 1×10^4). This cycle was repeated five times.

NMR measurements

Spectra of selectively ^{13}C′/^{15}N-labeled Ras proteins were recorded on a Bruker AMX-600 spectrometer, and the spectrum of [^{15}N]Tyr^{32}-Ras protein was recorded on a Bruker DMX-500 spectrometer at a probe temperature of 30 °C. The NMR data were processed and analyzed with the Felix (Biosym-MSI), the Azara (W. Boucher, unpublished), the NMRPipe (Delaglio et al., 1995), and the NMRVIEW (Johnson and Blevins, 1994) programs on an Indigo^2 workstation (SGI). The 2D ^{1}H–^{15}N HSQC (Bodenhausen and Ruben, 1980) spectra were acquired with 64 (t_1) × 1024 (t_2) complex points using the water flip-back scheme (Grzesiek and Bax, 1993). The 2D ^{1}H–^{15}N HNCO (Ikura et al., 1990) spectra were acquired with 32 (t_1) × 512 (t_2) complex points. For each FID, 256 transients for HSQC, and 1024 transients for HNCO were accumulated. By zero-filling followed by Fourier transformation, spectra of 1024 (^{1}H) × 128 (^{15}N) points for selectively ^{13}C′/^{15}N-labeled Ras proteins and of 1024 (^{1}H) × 512 (^{15}N) points for [^{15}N]Tyr^{32}-Ras protein, respectively, were obtained.

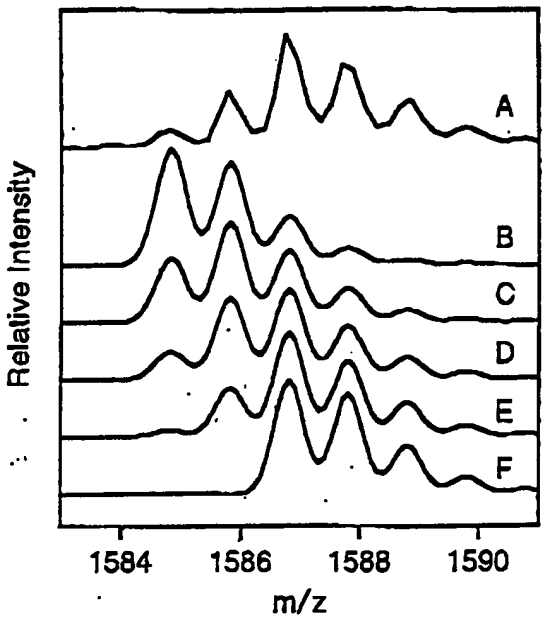


Figure 1. The isotopic distribution patterns for the Ras(136–148) fragment. (A) The observed pattern of the MALDI-TOF mass spectrum for the Ras(136–148) fragment of the [^{13}C′]Asp/[^{15}N]Ser-selectively labeled Ras protein. Simulated patterns for the Ras(136–148) fragment with various ^{15}N contents: 0% (B), 25% (C), 50% (D), 75% (E), and 100% (F).

Results and Discussion

Preparation of the selectively ^{13}C′/^{15}N-labeled Ras protein

For the [^{13}C′]Pro/[^{15}N]Thr-selective labeling of the Ras protein, a 15-mL reaction of the 'final' system was used with 2 mg each of DL-[^{13}C′]proline and L-[^{15}N]threonine. The labeled Ras protein was purified to homogeneity from the crude cell-free reaction mixture by successive chromatography on Mono Q and Superdex 75 columns. Finally, we obtained approximately 6 mg of the Ras protein, as determined by absorbance at 280 nm (data not shown).

Similarly, for the [^{13}C′]Asp/[^{15}N]Ser-selective labeling, a 15-mL reaction with 2 mg each of L-[^{13}C′]aspartate and L-[^{15}N]serine was used. We obtained about 6 mg of the Ras protein (data not shown).

Mass spectrum of the [^{13}C′]Asp/[^{15}N]Ser-selectively labeled Ras protein

To determine the level of ^{15}N enrichment of the [^{13}C′]Asp/[^{15}N]Ser-selectively labeled Ras protein, an aliquot of the sample was digested by trypsin, and was subjected to MALDI-TOF mass spectrometry. The isotopic distribution patterns of the peak for the fragment including two Ser residues and no Asp residue, Ras(136–148), were analyzed (Figure 1A). As the isotopic distribution pattern depends on the ^{15}N content, it is possible to estimate the actual ^{15}N content from the observed pattern by comparison with the simulated patterns from various ^{15}N contents (Figures 1B-F). The observed pattern from the labeled Ras(136-148) was most similar to the simulated pattern for the 75%-labeled Ras(136–148) (Figure 1E). Thus, the level of ^{15}N enrichment was estimated to be 75%. This enrichment level is higher than the case of the dual labeling by an *in vivo* system, in which the level of ^{15}N enrichment was less than 50% (Kainosho and Tsuji, 1982).

[^{13}C′]Pro/[^{15}N]Thr-selectively labeled Ras protein

To measure the ^{1}H–^{15}N HSQC and ^{1}H–^{15}N HNCO spectra (Figures 2A and B, respectively), 6 mg each of the labeled Ras•GMPPNP and the Raf RBD were used (the molar ratio was about 1 to 2). In the case of the Ras protein produced with L-[^{15}N]threonine by the *E. coli* expression system, the Gly and Ser residues were labeled in addition to the Thr residues, because the threonine aldolase catalyzes the threonine–glycine conversion and the serine hydroxymethyltransferase catalyzes the glycine–serine conversion (McIntosh and Dahlquist, 1990; Yamasaki et al., 1992). The number of observed cross peaks in the HSQC spectrum (Figure 2A) is equal to that of the Thr residues in the Ras protein, indicating that these metabolic conversions do not occur in the cell-free system. As the Pro–Thr sequence is unique in the Ras protein, the Thr^{35} cross peak was unambiguously identified by the HNCO experiment (Figure 2B).

[^{13}C′]Asp/[^{15}N]Ser-selectively labeled Ras protein

The ^{1}H–^{15}N HSQC and ^{1}H–^{15}N HNCO spectra (Figures 3A and B, respectively) were recorded using 6 mg each of the labeled Ras•GMPPNP and the Raf RBD (the molar ratio was about 1 to 2). Although the ^{15}N enrichment level of the sample is sufficiently high (about 75%), the signal-to-noise ratio of the HSQC spectrum was not as high as in the spectrum of the Ras protein by itself at the same Ras concentration (about 1 mM) (Muto et al., 1993; Ito et al., 1997). This is because of the long correlation time (approximately 21 ns; Terada et al., unpublished) of the Ras•Raf RBD complex (the total molecular mass is about 30 kDa). In the HSQC spectrum, seven peaks are observed, while there are eight serine residues in the Ras protein (Figure 3A). Since isotopic transfer from serine to any other amino acid could not occur in the cell-

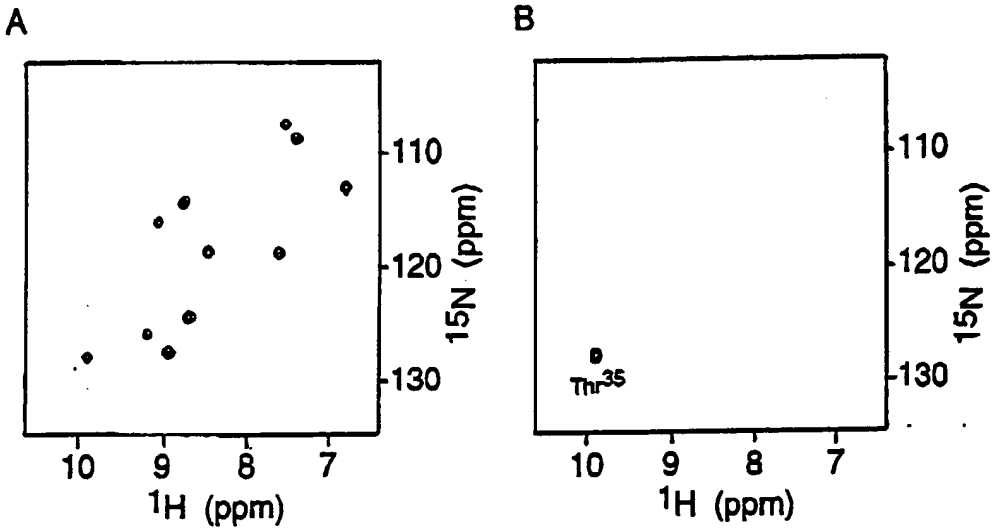


Figure 2. The HSQC spectrum (A) and the HNCO spectrum (B) of [^{13}C′]Pro/[^{15}N]Thr-selectively labeled Ras•GMPPNP bound to Raf RBD.

free system (Kigawa et al., 1995), we concluded that the seven peaks were derived from the Ser residues, and two of them were overlapped. There are two Asp–Ser sequences in the Ras protein: Asp^{38}–Ser^{39} and Asp^{105}–Ser^{106}, and two cross peaks were observed in the HNCO spectrum (Figure 3B). The peak at ^{1}HN = 7.51 ppm and ^{15}N = 110.64 ppm was assigned to Ser^{106}, by comparison with the assignment of Ser^{106} in both the GDP-bound Ras protein (Ras•GDP, Muto et al., 1993; Ito et al., 1997) and Ras•GMPPNP (Ito et al., 1997). This is also supported by the consideration that Ser^{106}is not located on the binding interface with Raf RBD, from the crystal structure of the complex of GMPPNP-bound Rap1A (Rap1A•GMPPNP), which is a homologue of Ras, and the Raf RBD (Nassar et al., 1995, 1996). The peak at ^{1}HN = 9.03 ppm and ^{15}N = 118.04 ppm was assigned to Ser^{39}.

Effector region of the Ras protein complexed with the Raf RBD

The backbone amide resonances of the residues in the effector region (including Thr^{35} and Ser^{39})and some other regions of Ras•GMPPNP were extremely broadened (Ito et al., 1997). This is because these regions slowly interconvert between two or more stable conformers ('regional polysterism', Ito et al., 1997). On the other hand, three cross peaks, Asp^{33}, Asp^{38}, and Asp^{57}, which were too broad to be detected for Ras•GMPPNP by itself (Ito et al., 1997), were clearly observed for Ras•GMPPNP in the complex with Raf RBD. Moreover, the Thr^{35} and Ser^{39} cross peaks could also be observed, as demonstrated in this study. Thus, the conformations of the effector region of Ras•GMPPNP are mostly fixed upon binding with Raf RBD.

The chemical shift differences of Thr^{35} and Ser^{39}between Ras•GMPPNP complexed with the Raf RBD and Ras•GDP (Muto et al., 1993; Ito et al., 1997) were large (Thr^{35}, Δ^{1}HN = 0.97 ppm and Δ^{15}N = 16.15 ppm; Ser^{39}, Δ^{1}HN = 0.67 ppm and Δ^{15}N = 2.40 ppm), while that of Ser^{106} was negligibly small (Δ^{1}HN = 0.07 ppm and Δ^{15}N = 0.08 ppm). This indicates that the environments of the Thr^{35} and Ser^{39} amide groups dramatically change upon binding with Raf RBD. In the tertiary structure from the X-ray crystallography study, the Thr^{35} amide group of Ras•GMPPNP forms a hydrogen bond with the γ-phosphate group of GMPPNP (Pai et al., 1989), while that of Ras•GDP does not interact with any atoms (Milburn et al., 1990). This is also the case with Rap1A•GMPPNP complexed with Raf RBD (Nassar et al., 1995, 1996). Our result suggests that the Thr^{35} of Ras•GMPPNP complexed with Raf RBD also forms a hydrogen bond with the γ-phosphate group of the GMPPNP. The orientation of Ser^{39} of the Ras protein is almost the same between the GDP-bound and GMPPNP-bound forms, according to the X-ray study (Pai et al., 1989; Milburn et al., 1990). In the Rap1A•GMPPNP structure in the complex with Raf RBD, the amide group of Ser^{39} of Rap1A forms a hydrogen bond with the carbonyl group of Arg^{67} of Raf RBD (Nassar et al., 1995, 1996). Our result suggests that, in Ras•GMPPNP complexed with Raf

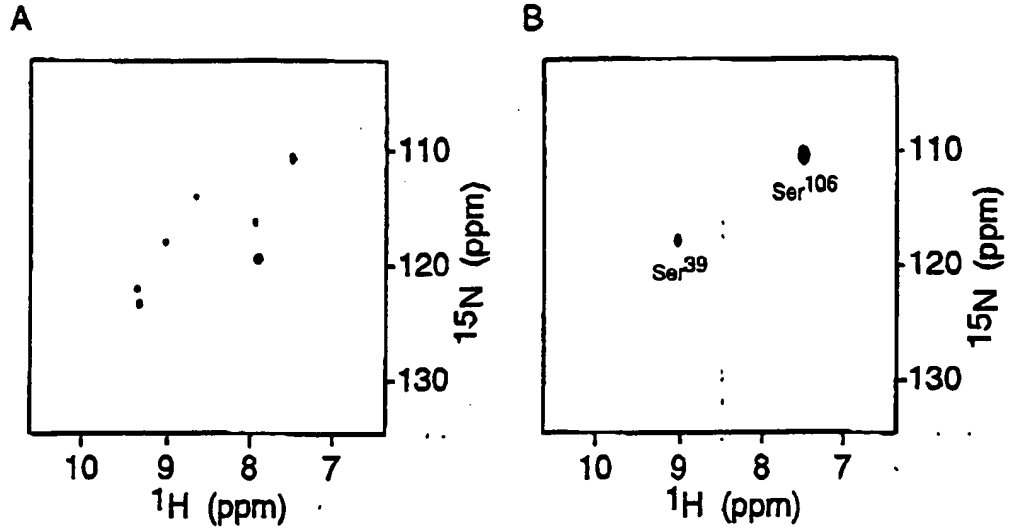


Figure 3. The HSQC spectrum (A) and the HNCO spectrum (B) of [^{13}C']Asp/[^{15}N]Ser-selectively labeled Ras•GMPPNP bound to Raf RBD.

RBD, Ser^{39} also forms a hydrogen bond with some residue (presumably Arg^{67}) of Raf RBD.

Scheme for the site-directed ^{15}N-labeling of Tyr^{32} of the Ras protein

Figure 4 summarizes the procedures for the site-directed stable-isotope amino acid labeling of the Ras protein. The codon for Tyr^{32} of the Ras protein was replaced by an amber termination codon (TAG) on the plasmid template (pK7-RasY32am). An *E. coli* tyrosine suppressor, tRNA$^{Tyr}_{CUA}$, was prepared by *in vitro* transcription with T7 RNA polymerase (25 mg purified tRNA$^{Tyr}_{CUA}$ per 10-mL transcription reaction mixture). The tRNA$^{Tyr}_{CUA}$ was precharged with L-[^{15}N]tyrosine by the *E. coli* TyrRS. The cell-free coupled transcription-translation reaction was carried out with [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$ and pK7-RasY32am as the template DNA; the amber codon (UAG) in the mRNA was recognized by the [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$ and then L-[^{15}N]tyrosine was site-specifically incorporated into Tyr^{32} of the Ras protein. The amber codon was competitively recognized by release factor 1 (RF1), and the abortive product, the Ras fragment consisting of residues 1–31, Ras(1–31), was also synthesized. The details of the translation experiments will be described below.

Time courses of cell-free protein synthesis

In order to optimize the reaction time for the site-directed isotope labeling (SDIL), the translation kinetics were analyzed on a small scale by the use of L-[^{14}C]tyrosine and L-[^{14}C]leucine, instead of L-[^{15}N]tyrosine and unlabeled L-leucine, respectively, as in the following. More than 50% of the suppressor tRNA$^{Tyr}_{CUA}$ was aminoacylated with L-[^{14}C]tyrosine, according to the small-scale preparation protocol (see Materials and Methods). The [^{14}C]Tyr-containing protein was efficiently synthesized with the SDIL system; the accumulated amount increased, and finally reached a plateau of about 100 μg/mL at about 45 min (Figure 5, crosses). This amount is approximately one-third of that synthesized in 45 min with the conventional synthesis system with no amber mutation (Figure 5, filled circles). Although the synthesis of the [^{14}C]Tyr-containing protein stopped at about 45 min, the synthesis of the [^{14}C]Leu-containing protein with the SDIL system still continued (Figure 5, open circles). At 45 min, the [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ appears to be exhausted (Figure 5, filled squares), and concomitantly the synthesis of the [^{14}C]Tyr-containing protein stopped. Therefore, the content of the [^{14}C]Tyr-containing protein was relatively high (about 70%) until 30 min, but then decreased because the unlabeled protein was synthesized by recharging of the deacylated suppressor tRNA$^{Tyr}_{CUA}$ with unlabeled L-tyrosine by the endogenous TyrRS.

SDS-PAGE analysis of the cell-free synthesized proteins

The proteins synthesized with [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ and/or L-[^{14}C]leucine by the small-scale reaction protocol were analyzed by SDS-PAGE with tricine buffer, followed by autoradiography (Figure 6). When the cell-free protein synthesis with pK7-RasY32am

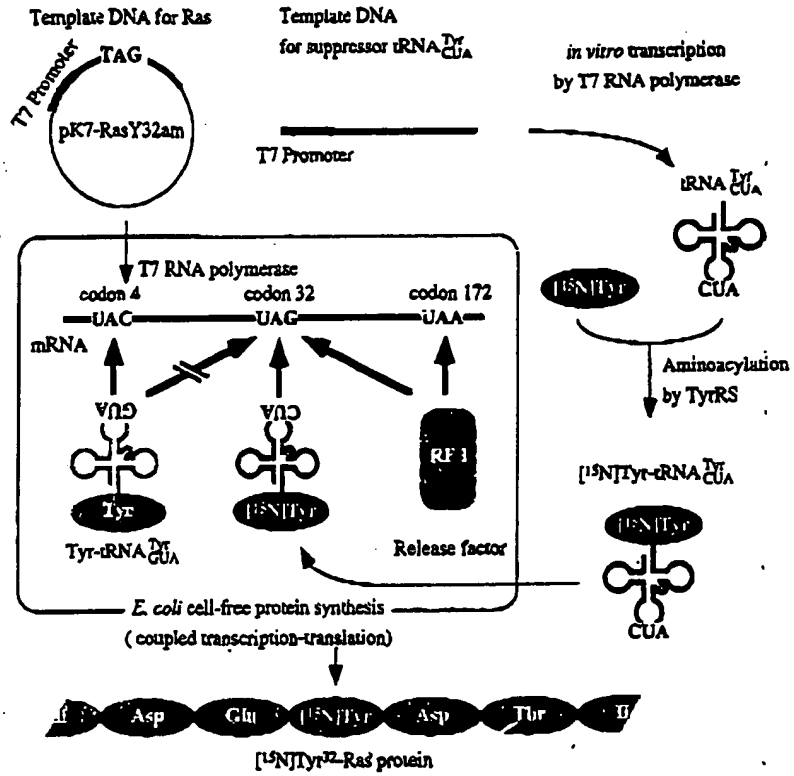


Figure 4. Site-directed stable-isotope labeling of the Ras protein.

lacked [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$, only the abortive product, Ras(1–31), was synthesized (Figure 6, lane 3). In contrast, the Ras protein was primarily produced in the presence of [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ (Figure 6, lane 4). The mobility of this Ras protein was confirmed to be equal to that of the authentic Ras protein, which was produced with the conventional synthesis system containing the original pK7-Ras plasmid without the amber mutation (Figure 6, lane 2). The suppression efficiency was estimated to be higher than 90% by analysis of the autoradiogram; [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ efficiently suppressed the amber mutation. It was shown that the Ras protein, as synthesized through amber suppression, definitely contained the [^{14}C]Tyr residue that had been precharged to the suppressor tRNA$^{Tyr}_{CUA}$ (Figure 6, lane 5).

Preparation of the [^{15}N]Tyr^{32}-Ras protein

The tRNA$^{Tyr}_{CUA}$ aminoacylated with L-[^{15}N]tyrosine was prepared according to the large-scale preparation protocol, as described in Materials and Methods. As the tRNA$^{Tyr}_{CUA}$ concentration was higher than that in the small-scale reaction protocol, we analyzed the aminoacylation efficiency from the L-[^{14}C]tyrosine incorporation, and confirmed that about 50% of the suppressor tRNA$^{Tyr}_{CUA}$ was aminoacylated (data not shown). The protein synthesis reaction was carried out with [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$. As much as 2.2 mg of the Ras protein could be purified from a 30-mL reaction. The small-scale analysis had predicted the production of about 3 mg of crude Ras protein per 30 mL reaction. Considering that the 2.2 mg were obtained after the protein purification, the present large-scale preparation of the site-specifically labeled Ras protein was quite successful. The SDS-PAGE analysis revealed that the purified Ras protein formed a single band, and its molecular mass was equal to that of the authentic Ras protein (Figure 7). This was also confirmed by mass spectrometry analysis (data not shown). This shows that the Ras protein was clearly separated from the abortive product, Ras(1–31), if any, by the chromatographic purification procedure.

Mass spectrum of the [^{15}N]Tyr^{32}-Ras protein

In order to estimate the ^{15}N content of the Tyr^{32} residue, labeled and unlabeled Ras protein samples were digested by *Achromobacter lyticus* protease I, and were subjected to MALDI-TOF mass spectrom-

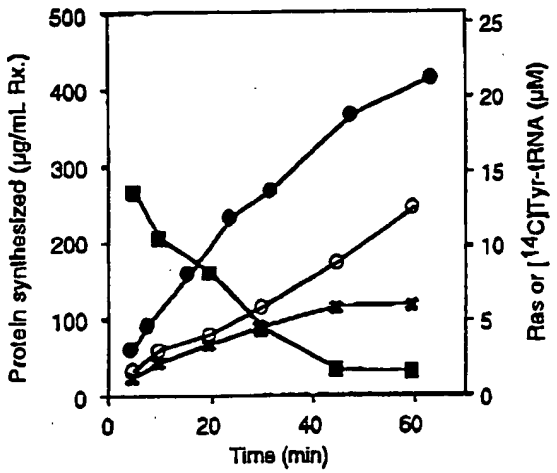


Figure 5. Time courses of cell-free protein synthesis. Each amount of synthesized protein at a given time is normalized to μg Ras protein per 1 mL of the reaction mixture (left-hand ordinate), and is also indicated as μM units (right-hand ordinate). The amount of [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ is normalized to μM. The [^{14}C]Tyr-containing protein (x) and the total protein (○) (see Materials and Methods) as synthesized by the cell-free protein synthesis system involving the amber-mutant pK7-RasY32am plasmid as the template together with the [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ (the SDIL system), the protein as synthesized by the conventional synthesis system with the pK7-Ras plasmid without the amber mutation (●), and the amount of [^{14}C]Tyr-tRNA$^{Tyr}_{CUA}$ in the SDIL system (■).

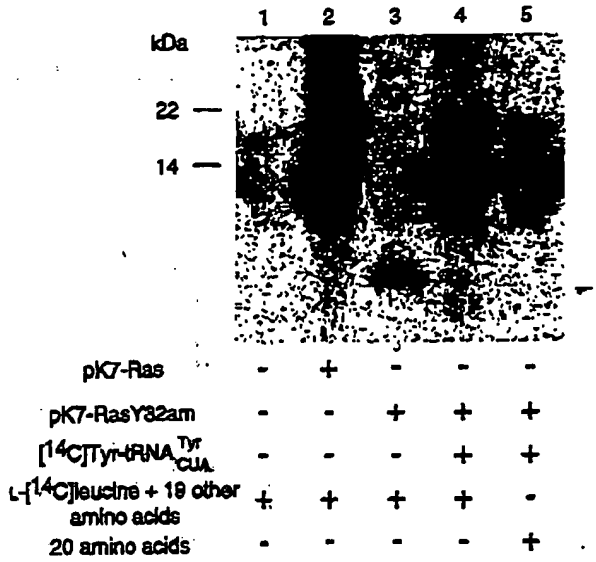


Figure 6. Tricine-SDS-PAGE analysis of the cell-free reaction products (autoradiography). The components in the reactions are indicated for each sample.

etry. The isotopic distribution patterns of the peak for the fragment including Tyr^{32}, Ras(17–42), were then analyzed (Figure 8A). The observed pattern from the labeled Ras protein was most similar to the simulated

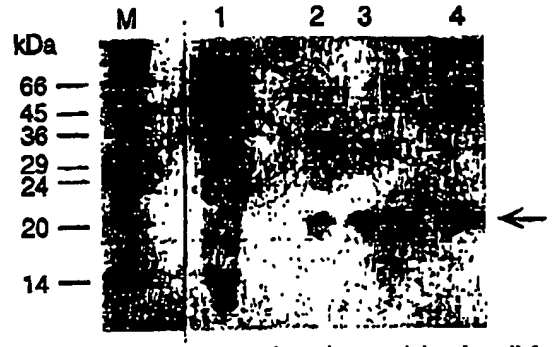


Figure 7. SDS-PAGE analysis of samples containing the cell-free synthesized Ras protein at each purification step (CBB staining). Lane 1: crude reaction product; lane 2: Resource Q column chromatography fraction; lane 3: Superdex 75 column chromatography fraction; lane 4: the authentic Ras protein that was overexpressed *in vivo* and purified; lane M: molecular mass markers.

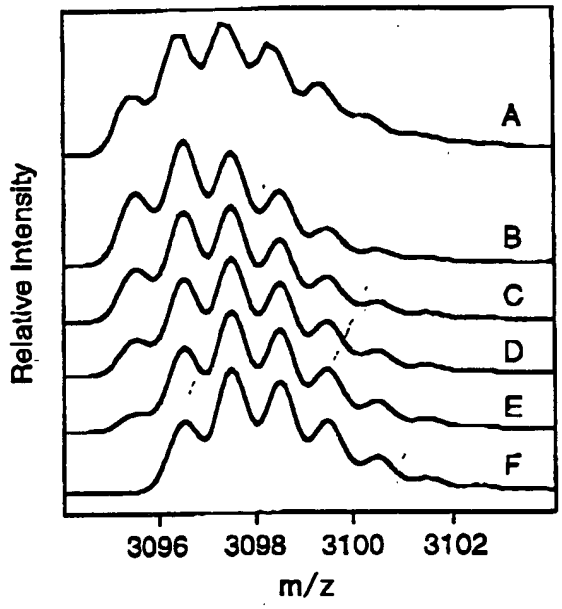


Figure 8. The isotopic distribution patterns for the Ras(17–42) fragment including Tyr^{32}. (A) The observed pattern of the MALDI-TOF mass spectrum for the Ras(17–42) fragment of the [^{15}N]Tyr^{32}-Ras protein. Simulated patterns for the Ras(17–42) fragment with various ^{15}N contents: 0% (B), 25% (C), 50% (D), 75% (E), and 100% (F).

pattern for the 50%-labeled Ras(17–42) (Figure 8D). Thus, the ^{15}N content was estimated to be about 50%, which is slightly lower than that predicted by the small-scale analysis described above (about 70%). We are planning to solve this problem by lowering the rate of non-specific deacylation of [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$, for example, by using overproduced elongation factor Tu. These efforts may also reduce the total amount of

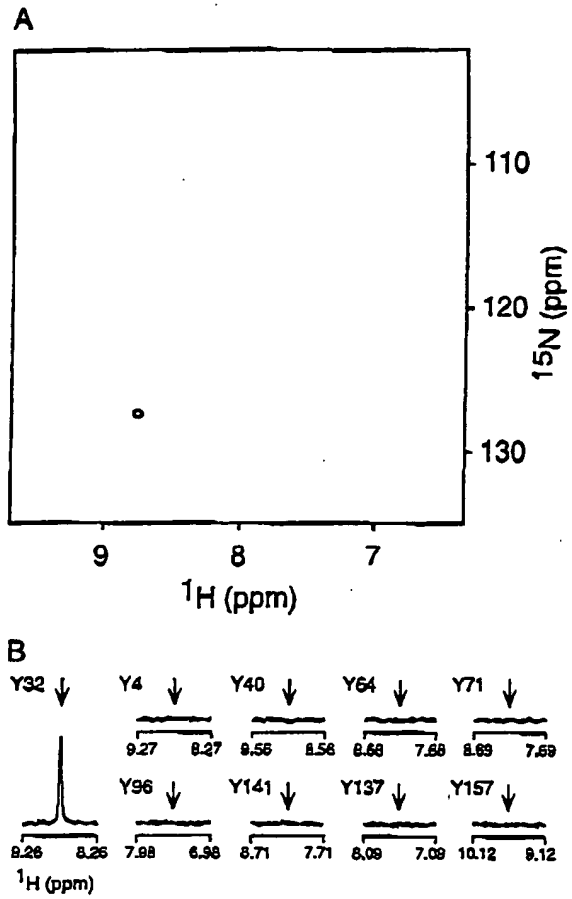


Figure 9. The HSQC spectrum of the [^{15}N]Tyr^{32}-Ras protein (A) and its cross sections (B). Arrows indicate the positions where we have observed, for the uniformly ^{15}N-labeled Ras protein, cross peaks of Tyr residues (Muto et al., 1993; Ito et al., 1997).

suppressor tRNA required for the site-directed labeling.

NMR spectrum of the [^{15}N]Tyr^{32}-Ras protein

The [^{15}N]Tyr^{32}-Ras protein (2.2 mg) in the GDP-bound form was subjected to NMR analysis. The HSQC spectrum has only one cross peak (Figure 9A). The ^{1}H and ^{15}N chemical shifts of the cross peak were equal to those for the Tyr^{32} residue of Ras•GDP (Muto et al., 1993; Ito et al., 1997). No cross peak corresponding to Tyr residues other than Tyr^{32} was observed in the spectrum (Figure 9B). Thus, Tyr^{32} was labeled site-specifically. Isotope scrambling of L-[^{15}N]tyrosine had been expected. However, even when a number of L-[^{15}N]tyrosine molecules were released from [^{15}N]Tyr-tRNA$^{Tyr}_{CUA}$ by hydrolysis prior to incorporation into the protein, they were not recycled by the endogenous TyrRS, because of dilution by the excess amount of unlabeled L-tyrosine in the cell-free protein synthesis system. Thus, we reduced the scrambling of [^{15}N]tyrosine to a negligible level.

Cell-free protein synthesis system for NMR structure analysis

The dual amino acid-selective ^{13}C–^{15}N labeling by cell-free protein synthesis is useful to decrease the number of cross peaks, and furthermore for reliable resonance assignment. The present study showed that these advantages are much enhanced by the use of cell-free protein synthesis as compared with the conventional recombinant techniques. The cell-free dual labeling is particularly powerful for structural investigation of a large protein or a protein complex focusing on a residue of functional importance.

In addition, we succeeded in obtaining a milligram quantity of a Ras protein that was isotopically labeled at a specific residue. The site-directed incorporation of a stable-isotope labeled amino acid other than tyrosine is possible if the aminoacylated suppressor tRNA is available. In this context, a number of suppressor tRNAs, which are effective *in vivo*, have been reported (Kleina et al., 1990; Normanly et al., 1990). Some of them show low aminoacylation activities *in vitro* (Shulman, 1991). For amino acids for which no efficient suppressor tRNA is available, the aminoacylated suppressor tRNA could be obtained by the chemical aminoacylation method (Ellman et al., 1991; Robertson et al., 1991; Lodder et al., 1997) or by aminoacylation with a mutant aminoacyl-tRNA synthetase engineered to aminoacylate the suppressor tRNA (Schmitt et al., 1993; Liu et al., 1997). Site-directed stable-isotope labeling drastically simplifies the observation and resonance-assignment procedures for a specified amino acid residue of particular interest, and will therefore be useful, for example, for analyzing local structures of large proteins and protein-protein interactions. Site-directed labeling will also be useful for some solid-state NMR methods, such as accurate analyses of the distance between two specified atoms (Peersen et al., 1992) and of the orientational constraint of a defined bond vector in an oriented sample (Ketchem et al., 1996). Thus, dual labeling and site-directed labeling by cell-free protein synthesis will be useful techniques for analyzing the structures of proteins.

Conclusions

Dual selective ^{13}C–^{15}N labeling of the Ras protein was achieved using cell-free protein synthesis. This allowed us to investigate the structure of the effector region of the Ras protein complexed with Raf RBD. We have also synthesized milligram quantities of a site-specifically stable-isotope labeled protein and have measured the ^{1}H–^{15}N HSQC spectrum of the protein. This method will be applicable to a wide range of solution and solid-state NMR studies.

Acknowledgements

We thank Dr. Y. Arata (Water Research Institute, Ibaraki, Japan) for encouragement. This work was supported in part by a grant for the Biodesign Research Program and a Special Grant for Promotion of Research, from the Institute of Physical and Chemical Research (RIKEN), Japan, a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture, Japan (08260222, 08558076, 08780570), a Grant-in-Aid (Bio Media Program) from the Ministry of Agriculture, Forestry and Fisheries, Japan (BMP 97-V-4-2), and a Grant-in-Aid ('Research for the Future' Program) from the Japan Society for the Promotion of Science, Japan (JSPS-RFTF 96100305).

References

Akashi, S., Shirouzu, M., Yokoyama, S. and Takio, K. (1996) *J. Mass Spectrom. Soc. Jpn.*, **44**, 269–277.

Arata, Y., Kato, K., Takahashi, H. and Shimada, I. (1994) *Methods Enzymol.*, **239**, 440–464.

Bain, J.D., Diala, E.S., Glabe, C.G., Wacker, D.A., Lyttle, M.H., Dix, T.A. and Chamberlin, A.R. (1991) *Biochemistry*, **30**, 5411–5421.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–199.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Ellman, J.A., Mendel, D., Anthony-Cahill, S., Noren, C.J. and Schultz, P.G. (1991) *Methods Enzymol.*, **202**, 301–336.

Ellman, J.A., Volkman, B.F., Mendel, D., Schultz, P.G. and Wemmer, D.E. (1992) *J. Am. Chem. Soc.*, **114**, 7959–7961.

Grzesick, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.

Ha, J.-M., Ito, Y., Kawai, G., Miyazawa, T., Miura, K., Ohtsuka, E., Noguchi, S., Nishimura, S. and Yokoyama, S. (1989) *Biochemistry*, **28**, 8411–8416.

Ikura, M., Kay, L.E. and Bax, A. (1990) *Biochemistry*, **29**, 4659–4667.

Ito, Y., Yamasaki, K., Iwahara, J., Terada, T., Kamiya, A., Shirouzu, M., Muto, Y., Kawai, G., Yokoyama, S., Laue, E.D., Wälchli, M., Shibata, T., Nishimura, S. and Miyazawa, T. (1997) *Biochemistry*, **36**, 9109–9119.

Johnson, B.A. and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.

Kainosho, M., Nagao, H. and Tsuji, T. (1987) *Biochemistry*, **26**, 1068–1075.

Kainosho, M., Nagao, H., Imamura, Y., Uchida, K., Tomonaga, N., Nakamura, Y. and Tsuji, T. (1985) *J. Mol. Struct.*, **126**, 549–562.

Kainosho, M. and Tsuji, T. (1982) *Biochemistry*, **21**, 6273–6279.

Kato, K., Matsunaga, C., Nishimura, Y., Wälchli, M., Kainosho, M. and Arata, Y. (1989) *J. Biochem. (Tokyo)*, **105**, 867–869.

Ketchem, R.R., Lee, K.C., Huo, S. and Cross, T.A. (1996) *J. Biomol. NMR*, **8**, 1–14.

Kigawa, T., Muto, Y. and Yokoyama, S. (1995) *J. Biomol. NMR*, **6**, 129–134.

Kim, D.-M., Kigawa, T., Choi, C.-Y. and Yokoyama, S. (1996) *Eur. J. Biochem.*, **239**, 881–886.

Kleina, L.G., Masson, J.M., Normanly, J., Abelson, J. and Miller, J.H. (1990) *J. Mol. Biol.*, **213**, 705–717.

Liu, D.R., Magliery, T.J., Pasternak, M. and Schultz, P.G. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 10092–10097.

Liu, X.M., Sonar, S., Lee, C.P., Coleman, M., RajBhandary, U.L. and Rothschild, K.J. (1995) *Biophys. Chem.*, **56**, 63–70.

Lodder, M., Golovine, S. and Hecht, S.M. (1997) *J. Org. Chem.*, **62**, 778–779.

McIntosh, L.P. and Dahlquist, F.W. (1990) *Q. Rev. Biophys.*, **23**, 1–38.

Milburn, M.V., Tong, L., deVos, A.M., Brunger, A., Yamaizumi, Z., Nishimura, S. and Kim, S.H. (1990) *Science*, **247**, 939–945.

Muto, Y., Yamasaki, K., Ito, Y., Yajima, S., Masaki, H., Uozumi, T., Wälchli, M., Nishimura, S., Miyazawa, T. and Yokoyama, S. (1993) *J. Biomol. NMR*, **3**, 165–184.

Nassar, N., Horn, G., Herrmann, C., Block, C., Janknecht, R. and Wittinghofer, A. (1996) *Nat. Struct. Biol.*, **3**, 723–729.

Nassar, N., Horn, G., Herrmann, C., Scherer, A., McCormick, F. and Wittinghofer, A. (1995) *Nature*, **375**, 554–560.

Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C. and Schultz, P.G. (1989) *Science*, **244**, 182–188.

Normanly, J., Kleina, L.G., Masson, J.M., Abelson, J. and Miller, J.H. (1990) *J. Mol. Biol.*, **213**, 719–726.

Pai, E.F., Kabsch, W., Krengel, U., Holmes, K.C., John, J. and Wittinghofer, A. (1989) *Nature*, **341**, 209–214.

Peersen, O.B., Yoshimura, S., Hojo, H., Aimoto, S. and Smith, S.O. (1992) *J. Am. Chem. Soc.*, **114**, 4332–4335.

Pratt, J.M. (1984) In *Transcription and Translation* (Eds, Hames, B.D. and Higgins, S.J.), IRL Press, Oxford, pp. 179–209.

Robertson, S.A., Ellman, J.A. and Schultz, P.G. (1991) *J. Am. Chem. Soc.*, **113**, 2722–2729.

Schagger, H. and von Jagow, G. (1987) *Anal. Biochem.*, **166**, 368–379.

Schmitt, E., Meinnel, T., Panvert, M., Mechulam, Y. and Blanquet, S. (1993) *J. Mol. Biol.*, **233**, 615–628.

Shulman, L.H. (1991) *Prog. Nucleic Acid Res. Mol. Biol.*, **41**, 23–87.

Sonar, S., Lee, C.-P., Coleman, M., Patel, N., Liu, X., Marti, T., Khorana, H.G., RajBhandary, U.L. and Rothschild, K.J. (1994) *Nat. Struct. Biol.*, **1**, 512–517.

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) *Methods Enzymol.*, **185**, 60–89.

Westler, W.M., Kainosho, M., Nagao, H., Tomonaga, N. and Markley, J.L. (1988a) *J. Am. Chem. Soc.*, **110**, 4093–4095.

Westler, W.M., Stockman, B.J., Markley, J.L., Hosoya, Y., Miyake, Y. and Kainosho, M. (1988b) *J. Am. Chem. Soc.*, **110**, 6256–6258.

Yamasaki, K., Muto, Y., Ito, Y., Wälchli, M., Miyazawa, T., Nishimura, S. and Yokoyama, S. (1992) *J. Biomol. NMR*, **2**, 71–82.

Zawadzki, V. and Gross, H.J. (1991) *Nucleic Acids Res.*, **19**, 1948.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 33/50, 33/542 — A1

(11) International Publication Number: WO 97/18471

(43) International Publication Date: 22 May 1997 (22.05.97)

(21) International Application Number: PCT/US96/18270

(22) International Filing Date: 13 November 1996 (13.11.96)

(30) Priority Data:
08/555,691 14 November 1995 (14.11.95) US

(71) Applicant: ABBOTT LABORATORIES [US/US]; CHAD 0377/AP6D-2, 100 Abbott Park Road, Abbott Park, IL 60064-3500 (US).

(72) Inventors: FESIK, Stephen, W.; 1099 Portsmouth Circle, Gurnee, IL 60031 (US). HAJDUK, Philip, J.; 693 Constitution #3, Palatine, IL 60074 (US).

(74) Agents: BROWDER, Monte, R. et al.; Abbott Laboratories, CHAD 0377/AP6D-2, 100 Abbott Park Road, Abbott Park, IL 60064-3500 (US).

(81) Designated States: AU, CA, IL, JP, MX, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: USE OF NUCLEAR MAGNETIC RESONANCE TO IDENTIFY LIGANDS TO TARGET BIOMOLECULES

(57) Abstract

The present invention provides a process for identifying compounds which bind to a specific target molecule. The process includes the steps of: a) generating a first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of a ^{15}N-labeled target molecule; b) exposing the labeled target molecule to one or a mixture of chemical compounds; c) generating a second two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of the labeled target molecule that has been exposed to one or a mixture of compounds in step (b); and d) comparing said first and second two-dimensional $^{15}N/^{1}H$ NMR correlation spectra to determine differences between said first and said second spectra, the differences indentifying the presence of one or more compounds that are ligands which have bound to the target molecule.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

Use of Nuclear Magnetic Resonance to Identify Ligands to Target Biomolecules

Technical Field of the Invention

The present invention pertains to a method for the screening of compounds for biological activity and to the determination of binding dissociation constants using two-dimensional $^{15}N/^{1}H$ NMR correlation spectral analysis to identify and design ligands that bind to a target biomolecule.

Background of the Invention

One of the most powerful tools for discovering new drug leads is random screening of synthetic chemical and natural product databases to discover compounds that bind to a particular target molecule (i.e., the identification of ligands of that target). Using this method, ligands may be identified by their ability to form a physical association with a target molecule or by their ability to alter a function of a target molecule.

When physical binding is sought, a target molecule is typically exposed to one or more compounds suspected of being ligands and assays are performed to determine if complexes between the target molecule and one or more of those compounds are formed. Such assays, as is well known in the art, test for gross changes in the target molecule (e.g., changes in size, charge, mobility) that indicate complex formation.

Where functional changes are measured, assay conditions are established that allow for measurement of a biological or chemical event related to the target molecule (e.g., enzyme catalyzed reaction, receptor-mediated enzyme activation). To identify an alteration, the function of the target molecule is determined before and after exposure to the test compounds.

Existing physical and functional assays have been used successfully to identify new drug leads for use in designing therapeutic compounds. There are, however, limitations inherent to those assays that compromise their accuracy, reliability and efficiency.

A major shortcoming of existing assays relates to the problem of "false positives". In a typical functional assay, a "false positive" is a compound that triggers the assay but which compound is not effective in eliciting the desired physiological response. In a typical physical assay, a "false positive" is a compound that, for example, attaches itself to the target but in a non-specific manner (e.g., non-specific binding). False positives are particularly prevalent and

problematic when screening higher concentrations of putative ligands because many compounds have non-specific affects at those concentrations.

In a similar fashion, existing assays are plagued by the problem of "false negatives", which result when a compound gives a negative response in the assay but which compound is actually a ligand for the target. False negatives typically occur in assays that use concentrations of test compounds that are either too high (resulting in toxicity) or too low relative to the binding or dissociation constant of the compound to the target.

Another major shortcoming of existing assays is the limited amount of information provided by the assay itself. While the assay may correctly identify compounds that attach to or elicit a response from the target molecule, those assays typically do not provide any information about either specific binding sites on the target molecule or structure activity relationships between the compound being tested and the target molecule. The inability to provide any such information is particularly problematic where the screening assay is being used to identify leads for further study.

It has recently been suggested that X-ray crystallography can be used to identify the binding sites of organic solvents on macromolecules. However, this method cannot determine the relative binding affinities at different sites on the target. It is only applicable to very stable target proteins that do not denature in the presence of high concentrations of organic solvents. Moreover, this approach is not a screening method for rapidly testing many compounds that are chemically diverse, but is limited to mapping the binding sites of only a few organic solvents due to the long time needed to determine the individual crystal structures.

Compounds are screened to identify leads that can be used in the design of new drugs that alter the function of the target biomolecule. Those new drugs can be structural analogs of identified leads or can be conjugates of one or more such lead compounds. Because of the problems inherent to existing screening methods, those methods are often of little help in designing new drugs.

There continues to be a need to provide new, rapid, efficient, accurate and reliable means of screening compounds to identify and design ligands that specifically bind to a particular target.

Brief Summary of the Invention

In one aspect, the present invention provides a process of screening compounds for biological activity to identify ligands that bind to a specific target molecule. That process comprises the steps of: a) generating a first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of a ^{15}N-labeled target molecule; b) exposing the labeled target molecule to one or a mixture of chemical compounds; c) generating a second two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of the labeled target molecule that has been exposed to one or a mixture of compounds in step (b); and d) comparing said first and second two-dimensional $^{15}N/^{1}H$ NMR correlation spectra to determine differences between said first and said second spectra, the differences identifying the presence of one or more compounds that are ligands which have bound to the target molecule.

Where the process of the present invention screens more than one compound in step (b), that is, a mixture of compounds, and where a difference between the first spectrum generated from the target molecule alone and that generated from the target molecule in the presence of the mixture, additional steps are performed to identify which specific compound or compounds contained in the mixture is binding to the target molecule. Those additional steps comprise the steps of e) exposing the ^{15}N-labeled target molecule individually to each compound of the mixture, f) generating a two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of the labeled target molecule that has been individually exposed to each compound; and g) comparing each spectrum generated in step f) to the first spectrum generated from the target molecule alone to determine differences in any of those compared spectra, the differences identifying the presence of a compound that is a ligand which has bound to the target molecule.

Because the chemical shift values of the particular $^{15}N/^{1}H$ signals in the two-dimensional correlation spectrum correspond to known specific locations of atomic groupings in the target molecule (e.g., the N-H atoms of the amide or peptide link of a particular amino acid residue in a polypeptide), the process of the present invention allows not only for the for identification of which compound(s) bind to a particular target molecule, but also permit the determination of the particular binding site of the ligand on the target molecule.

In a second aspect, the present invention provides a process of determining the dissociation constant, K_D, for a given ligand and its target molecule. That process comprises the steps of a) generating a first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of a ^{15}N-labeled target molecule; b) exposing the labeled

target molecule to various concentrations of a ligand; c) generating a two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum at each concentration of ligand in step (b); d) comparing each spectrum from step (c) to the first spectrum from step (a); and e) calculating the dissociation constant between the target molecule and the ligand from those differences according to the equation:

$$K_D = \frac{([P]_o - x)([L]_o - x)}{x}$$

An advantageous aspect of the present invention is the capability of the process of the present invention to determine the dissociation constant of one ligand of the target molecule in the presence of a second molecule already bound to the ligand. This is generally not possible with prior art methods which employ "wet chemical" analytical methods of determining binding of a ligand to a target molecule substrate.

In this preferred embodiment, the process of determining the dissociation constant of a ligand can be performed in the presence of a second bound ligand. In accordance with this embodiment, the ^{15}N-labeled target molecule is bound to that second ligand before exposing that target to the test compounds.

The ability of the present method to determine not only the existence of binding between one ligand and the target molecule, but also the particular site of binding in the presence of a second bound ligand permits the capability to design a drug that comprises two or more linked moieties made up of the ligands.

This method uses the two-dimensional $^{15}N/^{1}H$ NMR correlation spectroscopic screening process as set forth above to identify a first and subsequent ligands that bind to the target molecule. A complex of the target molecule and two or more ligands is formed and the three-dimensional structure of that complex is determined preferably using NMR spectroscopy or X-ray crystallography. That three-dimensional structure is used to determine the spatial orientation of the ligands relative to each other and to the target molecule.

Based on the spatial orientation, the ligands are linked together to form the drug. The selection of an appropriate linking group is made by maintaining the spatial orientation of the ligands to one another and to the target molecule based upon principles of bond angle and bond length information well known in the organic chemical art.

Thus, the molecular design method comprises identifying a first ligand moiety to the target molecule using two-dimensional $^{15}N/^{1}H$ NMR correlation spectroscopy; identifying subsequent ligand moieties to the target molecule using two-dimensional $^{15}N/^{1}H$ NMR correlation spectroscopy; forming a complex of the first and subsequent ligand moieties to the target molecule; determining the three dimensional structure of the complex and, thus, the spatial orientation of the first and subsequent ligand moieties on the target molecule; and linking the first and subsequent ligand moieties to form the drug to maintain the spatial orientation of the ligand moieties.

The identification of subsequent ligand moieities can be performed in the absence or presence of the first ligand (e.g., the target molecule can be bound to the first ligand before being exposed to the test compounds for identification of the second ligand).

In a preferred embodiment, the target molecule used in a screening or design process is a polypeptide. The polypeptide target is preferably produced in recombinant form from a host cell transformed with an expression vector that contains a polynucleotide that encodes the polypeptide, by culturing the transformed host cell in a medium that contains an assimilable source of ^{15}N such that the recombinantly produced polypeptide is labeled with ^{15}N.

Brief Description of the Drawings

In the drawings which form a portion of the specification:

FIG. 1 shows a $^{15}N/^{1}H$ correlation spectrum of the DNA binding domain of uniformly ^{15}N-labeled human papillomavirus E2. The spectrum (80 complex points, 4 scans/fid) was acquired on a 0.5 mM sample of E2 in 20 mM phosphate (pH 6.5), 10 mM dithiothreitol (DTT) and 10% deuterium oxide (D_2O).

FIG. 2 shows $^{15}N/^{1}H$ correlation spectra of the DNA binding domain of uniformly ^{15}N-labeled human papillomavirus E2 before (thin multiple contours) and after (thick single contours) addition of a final test compound. The final concentration of compound was 1.0 mM. All other conditions are as stated in FIG. 1. Selected residues that show significant changes upon binding are indicated.

FIG. 3 shows $^{15}N/^{1}H$ correlation spectra of the DNA binding domain of uniformly ^{15}N-labeled human papillomavirus E2 before (thin multiple contours) and after (thick single contours) addition of a second test compound. The final concentration of compound was 1.0 mM. All other conditions are as stated in FIG. 1. Selected residues that show significant changes upon binding are indicated.

FIG. 4 shows $^{15}N/^{1}H$ correlation spectra of the catalytic domain of uniformly ^{15}N-labeled stromelysin before (thin multiple contours) and after (thick single contours) addition of a test compound. The final concentration of compound was 1.0 mM. The spectra (80 complex points, 8 scans/fid) were acquired on a 0.3 mM sample of SCD in 20 mM TRIS (pH 7.0), 20 mM $CaCl_2$ and 10% D_2O. Selected residues that show significant changes upon binding are indicated.

FIG. 5 shows $^{15}N/^{1}H$ correlation spectra of the Ras-binding domain of uniformly ^{15}N-labeled RAF peptide (residues 55-132) before (thin multiple contours) and after (thick single contours) addition of a test compound. The final concentration of compound was 1.0 mM. The spectra (80 complex points, 8 scans/fid) were acquired on a 0.3 mM sample of the RAF fragment in 20 mM phosphate (pH 7.0), 10 mM DTT and 10% D_2O. Selected residues that show significant changes upon binding are indicated.

FIG. 6 shows $^{15}N/^{1}H$ correlation spectra of uniformly ^{15}N-labeled FKBP before (thin multiple contours) and after (thick single contours) addition of a test compound. The final concentration of compound was 1.0 mM. The spectra (80 complex points, 4 scans/fid) was acquired on a 0.3 mM sample of FKBP in 50 mM phosphate (pH 6.5), 100 mM NaCl and 10% D_2O. Selected residues that show significant changes upon binding are indicated.

FIG. 7 shows a first depiction of the NMR-derived structure of the DNA-binding domain of E2. The two monomers of the symmetric dimer are oriented in a top-bottom fashion, and the N- and C-termini of each monomer are indicated (N and C for one monomer, N* and C* for the other). Shown in ribbons are the residues which exhibit significant chemical shift changes ($\Delta\delta(^{1}H)>0.04$ ppm; $\Delta\delta(^{15}N)>0.1$ ppm) upon binding to a first test compound. These residues correspond to the DNA-recognition helix of E2. Selected residues are numbered for aid in visualization.

FIG. 8 shows a second depiction of the NMR-derived structure of the DNA-binding domain of E2. The two monomers of the symmetric dimer are oriented in a top-bottom fashion, and the N- and C-termini of each monomer are indicated (N and C for one monomer, N* and C* for the other). Shown in ribbons are the residues which exhibit significant chemical shift changes ($\Delta\delta(^{1}H)$>0.04 ppm; $\Delta\delta(^{15}N)$ >0.1 ppm) upon binding to a second test compound. These residues are located primarily in the dimer interface region. Selected residues are numbered for aid in visualization.

FIG. 9 shows a depiction of the NMR-derived structure of the catalytic domain of stromelysin. The N- and C-termini are indicated. Shown in ribbons are the residues which exhibit significant chemical shift changes ($\Delta\delta(^{1}H)$>0.04 ppm; $\Delta\delta(^{15}N)$ >0.1 ppm) upon binding to a test compound. These either form part of the S1' binding site or are spatially proximal to this site. Selected residues are numbered for aid in visualization.

FIG. 10 shows a ribbon plot of a ternary complex of first and second ligands bound to the catalytic domain of stromelysin.

Detailed Description of the Invention

The present invention provides a rapid and efficient screening method for identifying ligands that bind to therapeutic target molecules.

Ligands are identified by testing the binding of molecules to a target molecule (e.g., protein, nucleic acid, etc.) by following, with nuclear magnetic resonance (NMR) spectroscopy, the changes in chemical shifts of the target molecule upon the addition of the ligand compounds in the database.

From an analysis of the chemical shift changes of the target molecule as a function of ligand concentration, the binding affinities of ligands for biomolecules are also determined.

The location of the binding site for each ligand is determined from an analysis of the chemical shifts of the biomolecule that change upon the addition of the ligand and from nuclear Overhauser effects (NOEs) between the ligand and biomolecule.

Information about the structure/activity relationships between ligands identified by such a process can then be used to design new drugs that serve as ligands to the target molecule. By way of example, where two or more ligands to a given target molecule are identified, a complex of those ligands and the target

molecule is formed. The spatial orientation of the ligands to each other as well as to the target molecule is derived from the three-dimensional structure. That spatial orientation defines the distance between the binding sites of the two ligands and the orientation of each ligand to those sites.

Using that spatial orientation data, the two or more ligands are then linked together to form a new ligand. Linking is accomplished in a manner that maintains the spatial orientation of the ligands to one another and to the target molecule.

There are numerous advantages to the NMR-based discovery process of the present invention. First, because a process of the present invention identifies ligands by directly measuring binding to the target molecule, the problem of false positives is significantly reduced. Because the present process identifies specific binding sites to the target molecule, the problem of false positives resulting from the non-specific binding of compounds to the target molecule at high concentrations is eliminated.

Second, the problem of false negatives is significantly reduced because the present process can identify compounds that specifically bind to the target molecule with a wide range of dissociation constants. The dissociation or binding constant for compounds can actually be determined with the present process.

Other advantages of the present invention result from the variety and detailed data provided about each ligand from the discovery process.

Because the location of the bound ligand can be determined from an analysis of the chemical shifts of the target molecule that change upon the addition of the ligand and from nuclear Overhauser effects (NOEs) between the ligand and biomolecule, the binding of a second ligand can be measured in the presence of a first ligand that is already bound to the target. The ability to simultaneously identify binding sites of different ligands allows a skilled artisan to 1) define negative and positive cooperative binding between ligands and 2) design new drugs by linking two or more ligands into a single compound while maintaining a proper orientation of the ligands to one another and to their binding sites.

Further, if multiple binding sites exist, the relative affinity of individual binding moieties for the different binding sites can be measured from an analysis of the chemical shift changes of the target molecule as a function of the added concentration of the ligand. By simultaneously screening numerous structural analogs of a given compound, detailed structure/activity relationships about ligands is provided.

In its principal aspect, the present invention provides a process of screening compounds to identify ligands that bind to a specific target molecule. That process comprises the steps of: a) generating a first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of a ^{15}N-labeled target molecule; b) exposing the labeled target molecule to one or more compounds; c) generating a second two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of the labeled target molecule that has been exposed to the compounds of step (b); and d) comparing the first and second spectra to determine whether differences in those two spectra exist, which differences indicate the presence of one or more ligands that have bound to the target molecule.

Where a process of the present invention screens more than one compound in step (b) and where a difference between spectra is observed, additional steps are performed to identify which specific compound is binding to the target molecules. Those additional steps comprise generating a two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum for each individual compound and comparing each spectrum to the first spectrum to determine whether differences in any of those compared spectra exist, which differences indicate the presence of a ligand that has bound to the target molecule.

Any ^{15}N-labeled target molecule can be used in a process of the present invention. Because of the importance of proteins in medicinal chemistry, a preferred target molecule is a polypeptide. The target molecule can be labeled with ^{15}N using any means well known in the art. In a preferred embodiment, the target molecule is prepared in recombinant form using transformed host cells. In an especially preferred embodiment, the target molecule is a polypeptide. Any polypeptide that gives a high resolultion NMR spectrum and can be partially or uniformly labeled with ^{15}N can be used. The preparation of uniformly ^{15}N-labeled exemplary polypeptide target molecules is set forth hereinafter in the Examples.

A preferred means of preparing adequate quantities of uniformly ^{15}N-labeled polypeptides is to transform a host cell with an expression vector that contains a polynucleotide that encodes that polypeptide and culture the transformed cell in a culture medium that contains assimilable sources of ^{15}N. Assimilable sources of ^{15}N are well known in the art. A preferred such source is $^{15}NH_4Cl$.

Means for preparing expression vectors that contain polynucleotides encoding specific polypeptides are well known in the art. In a similar manner, means for transforming host cells with those vectors and means for culturing those

transformed cells so that the polypeptide is expressed are also well known in the art.

The screening process of the present invention begins with the generation or acquisition of a two-dimensional $^{15}N/^{1}H$ correlation spectrum of the labeled target molecule. Means for generating two-dimensional $^{15}N/^{1}H$ correlation spectra are well known in the art (see, e.g., D. A. Egan et al., Biochemistry, 32(8): 1920-1927 (1993); Bax, A., Grzesiek, S., Acc. Chem. Res., 26(4): 131-138 (1993)).

The NMR spectra that are typically recorded in the screening procedure of the present invention are two-dimensional $^{15}N/^{1}H$ heteronuclear single quantum correlation (HSQC) spectra. Because the $^{15}N/^{1}H$ signals corresponding to the backbone amides of the proteins are usually well-resolved, the chemical shift changes for the individual amides are readily monitored.

In generating such spectra, the large water signal is suppressed by spoiling gradients. To facilitate the acquisition of NMR data on a large number of compounds (e.g., a database of synthetic or naturally occurring small organic compounds), a sample changer is employed. Using the sample changer, a total of 60 samples can be run unattended. Thus, using the typical acquisition parameters (4 scans per free induction decay (fid), 100-120 HSQC spectra can be acquired in a 24 hour period.

To facilitate processing of the NMR data, computer programs are used to transfer and automatically process the multiple two-dimensional NMR data sets, including a routine to automatically phase the two-dimensional NMR data. The analysis of the data can be facilitated by formatting the data so that the individual HSQC spectra are rapidly viewed and compared to the HSQC spectrum of the control sample containing only the vehicle for the added compound (DMSO), but no added compound. Detailed descriptions of means of generating such two-dimensional $^{15}N/^{1}H$ correlation spectra are set forth hereinafter in the Examples.

A representative two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of an ^{15}N-labeled target molecule (polypeptide) is shown in FIG. 1 (the DNA-binding domain of the E2 protein).

Following acquisition of the first spectrum, the labeled target molecule is exposed to one or more test compounds. Where more than one test compound is to be tested simultaneously, it is preferred to use a database of compounds such as a plurality of small molecules. Such molecules are typically dissolved in

perdeuterated dimethylsulfoxide . The compounds in the database can be purchased from vendors or created according to desired needs.

Individual compounds can be selected *inter alia* on the basis of size (molecular weight = 100-300) and molecular diversity. Compounds in the collection can have different shapes (e.g., flat aromatic rings(s), puckered aliphatic rings(s), straight and branched chain aliphatics with single, double, or triple bonds) and diverse functional groups (e.g., carboxylic acids, esters, ethers, amines, aldehydes, ketones, and various heterocyclic rings) for maximizing the possibility of discovering compounds that interact with widely diverse binding sites.

The NMR screening process of the present invention utilizes ligand concentrations ranging from about 0.1 to about 10.0 mM. At these concentrations, compounds which are acidic or basic can significantly change the pH of buffered protein solutions. Chemical shifts are sensitive to pH changes as well as direct binding interactions, and "false positive" chemical shift changes, which are not the result of ligand binding but of changes in pH, can therefore be observed. It is thus necessary to ensure that the pH of the buffered solution does not change upon addition of the ligand. One means of controlling pH is set forth below.

Compounds are stored at 263°K as 1.0 and 0.1 M stock solutions in dimethylsulfoxide (DMSO). This is necessary because of the limited solubility of the ligands in aqueous solution. It is not possible to directly adjust the pH of the DMSO solution. In addition, HCl and NaOH form insoluble salts in DMSO, so alternative acids and bases must be used. The following approach has been found to result in stable pH.

The 1.0 M stock solutions in DMSO are diluted 1:10 in 50 mM phosphate, pH 7.0. The pH of that diluted aliquot solution is measured. If the pH of the aliquot is unchanged (i.e., remains at 7.0), a working solution is made by diluting the DMSO stock solution 1:10 to make a 0.1 M solution and that solution is stored.

If the pH of the diluted aliquot is less than 7.0, ethanolamine is added to the 1.0 M stock DMSO solution, that stock solution is then diluted 1:10 with phosphate buffer to make another aliquot, and the pH of the aliquot rechecked.

If the pH of the diluted aliquot is greater than 7.0, acetic acid is added to the 1.0 M stock DMSO solution, that stock solution is then diluted 1:10 with phosphate buffer to make another aliquot, and the pH of the aliquot rechecked.

Ethanolamine and acetic acid are soluble in DMSO, and the proper equivalents are added to ensure that upon transfer to aqueous buffer, the pH is unchanged. Adjusting the pH is an interactive process, repeated until the desired result is obtained.

Note that this procedure is performed on 1:10 dilutions of 1.0 M stock solutions (100 mM ligand) to ensure that no pH changes are observed at the lower concentrations used in the experiments (0.1 to 10 mM) or in different/weaker buffer systems.

Following exposure of the ^{15}N-labeled target molecule to one or more test compounds, a second two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum is generated. That second spectrum is generated in the same manner as set forth above. The first and second spectra are then compared to determine whether there are any differences between the two spectra. Differences in the two-dimensional $^{15}N/^{1}H$ NMR correlation spectra that indicate the presence of a ligand correspond to ^{15}N-labeled sites in the target molecule. Those differences are determined using standard procedures well known in the art.

By way of example, FIGs. 2, 3, 4, 5 and 6 show comparisons of correlation spectra before and after exposure of various target molecules to various test compounds. A detailed description of how these studies were performed can be found hereinafter in Examples 2 and 3.

Particular signals in a two-dimensional $^{15}N/^{1}H$ correlation spectrum correspond to specific nitrogen and proton atoms in the target molecule (e.g., particular amides of the amino acid residues in the protein). By way of example, it can be seen from FIG. 2 that chemical shifts in a two-dimensional $^{15}N/^{1}H$ correlation of the DNA-binding domain of E2 exposed to a test compound occurred at residue positions 15 (I15), 21 (Y21), 22 (R22) and 23 (L23).

It can be seen from FIG. 2 that the binding of the ligand involved the isoleucine (Ile) residue at position 15, the tyrosine (Tyr) residue at position 21, the arginine (Arg) residue at position 22 and the leucine (Leu) residue at position 23. Thus, a process of the present invention can also be used to identify the specific binding site between a ligand and target molecule.

The region of the protein that is responsible for binding to the individual compounds is identified from the particular amide signals that change upon the addition of the compounds. These signals are assigned to the individual amide groups of the protein by standard procedures using a variety of well-established heteronuclear multi-dimensional NMR experiments.

To discover molecules that bind more tightly to the protein, molecules are selected for testing based on the structure/activity relationships from the initial screen and/or structural information on the initial leads when bound to the protein. By way of example, the initial screening may result in the identification of ligands, all of which contain an aromatic ring. The second round of screening would then use other aromatic molecules as the test compounds.

As set forth hereinafter in Example 2, an initial screening assay for binding to the catalytic domain of stromelysin identified two biaryl compounds as ligands. The second round of screening thus used a series of biaryl derivatives as the test compounds.

The second set of test compounds are initially screened at a concentration of 1 mM, and binding constants are measured for those that show affinity. Best leads that bind to the protein are then compared to the results obtained in a functional assay. Those compounds that are suitable leads are chemically modified to produce analogs with the goal of discovering a new pharmaceutical agent.

In another aspect, the present invention provides a process for determining the dissociation constant between a target molecule and a ligand that binds to that target molecule. That process comprises the steps of: a) generating a first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of a ^{15}N-labeled target molecule; b) titrating the labeled target molecule with various concentrations of a ligand; c) generating a two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum at each concentration of ligand from step (b); d) comparing each spectrum from step (c) both to the first spectrum from step (a) and to all other spectra from step (c) to quantify differences in those spectra as a function of changes in ligand concentration; and e) calculating the dissociation constant (K_D) between the target molecule and the ligand from those differences.

Because of their importance in medicinal chemistry, a preferred target molecule for use in such a process is a polypeptide. In one preferred embodiment, a process of determining the dissociation constant of a ligand can be performed in the presence of a second ligand. In accordance with this embodiment, the ^{15}N-labeled target molecule is bound to that second ligand before exposing that target to the test compounds.

Binding or dissociation constants are measured by following the $^{15}N/^{1}H$ chemical shifts of the protein as a function of ligand concentration. A known concentration ($[P]_0$) of the target moleule is mixed with a known concentration ($[L]_0$) of a previously identified ligand and the two-dimensional $^{15}N/^{1}H$

correlation spectrum was acquired. From this spectrum, observed chemical shift values (δ_{obs}) are obtained. The process is repeated for varying concentrations of the ligand to the point of saturation of the target molecule, when possible, in which case the limiting chemical shift value for saturation (δ_{sat}) is measured.

In those situations where saturation of the target molecule is achieved, the dissociation constant for the binding of a particular ligand to the targer molecule is calculated using the formula:

$$K_D = \frac{([P]_0 - x)([L]_0 - x)}{x}$$

where $[P]_0$ is the total molar concentration of target molecule; $[L]_0$ is the total molar concentration of ligand; and x is the molar concentration of the bound species. The value of x is determined from the equation:

$$x = \frac{\delta_{obs} - \delta_{free}}{\Delta}$$

where δ_{free} is the chemical shift of the free species; δ_{obs} is the observed chemical shift; and Δ is the differenec between the limiting chemical shift value for saturation (δ_{sat}) and the chemical shift value of the target molecule free of ligand (δ_{free}).

The dissociation constant is then determined by varying its value until a best fit to the observed data is obtained using standard curve-fitting statistical methods. In the case where δ_{sat} is not directly known, both K_D and δ_{sat} are varied and subjected to the same curve-fitting procedure.

The use of the process of the present invention to determine the dissociation or binding affinity of various ligands to various target molecules is set forth hereinafter in Examples 2 and 3.

Preferred target molecules, means for generating spectra, and means for comparing spectra are the same as set forth above.

The initial step in the design process is the identification of two or more ligands that bind to the specific target molecule. The identification of such ligands is done using two-dimensional $^{15}N/^{1}H$ NMR correlation spectroscopy as set forth above.

Once two or more ligands are identified as binding to the target molecule at different sites, a complex between the target molecule and ligands is formed. Where there are two ligands, that complex is a ternary complex. Quaternary and other complexes are formed where there are three or more ligands.

Complexes are formed by mixing the target molecule simultaneously or sequentially with the various ligands under circumstances that allow those ligands to bind the target. Means for determining those conditions are well known in the art.

Once that complex is formed, its three-dimensional structure is determined. Any means of determining three-dimensional structure can be used. Such methods are well known in the art. Exemplary and preferred methods are NMR and X-ray crystallography. The use of three-dimensional double- and triple resonance NMR to determine the three-dimensional structure of two ligands bound to the catalytic domain of stromelysin is set forth in detail hereinafter in Example 4.

An analysis of the three-dimensional structure reveals the spatial orientation of the ligands relative to each other as well as to the conformation of the target molecule. First, the spatial orientation of each ligand to the target molecule allows for identification of those portions of the ligand directly involved in binding (i.e., those portions interacting with the target binding site) and those portions of each ligand that project away from the binding site and which portions can be used in subsequent linking procedures.

Second, the spatial orientation data is used to map the positions of each ligand relative to each other. In other words, discrete distances between the spatially oriented ligands can be calculated.

Third, the spatial orientation data also defines the three-dimensional relationships amongst the ligands and the target. Thus, in addition to calculating the absolute distances between ligands, the angular orientations of those ligands can also be determined.

Knowledge of the spatial orientations of the ligands and target is then used to select linkers to link two or more ligands together into a single entity that contains all of the ligands. The design of the linkers is based on the distances and angular orientation needed to maintain each of the ligand portions of the single entity in proper orientation to the target.

The three-dimensional conformation of suitable linkers is well known or readily ascertainable by one of ordinary skill in the art. While it is theoretically possible to link two or more ligands together over any range of distance and three-dimensional projection, in practice certain limitations of distance and projection are preferred. In a preferred embodiment, ligands are separated by a distance of less than about 15 Angstroms (Å), more preferably less than about 10 Å and, even more preferably less than about 5 Å.

Once a suitable linker group is identified, the ligands are linked with that linker. Means for linking ligands are well known in the art and depend upon the chemical structure of the ligand and the linking group itself. Ligands are linked to one another using those portions of the ligand not directly involved in binding to the target molecule.

A detailed description of the design of a drug that inhibits the proteolytic activity of stromelysin, which drug was designed using a process of the present invention is set forth hereinafter in Example 4.

The following Examples illustrate preferred embodiments of the present invention and are not limiting of the specification and claims in any way.

Example 1

Preparation Of Uniformly ^{15}N-Labeled Target Molecules

A. Stromelysin

Human stromelysin is a 447-amino acid protein believed to be involved in proteolytic degradation of cartilage. Cartilage proteolysis is believed to result in degradative loss of joint cartilage and the resulting impairment of joint function observed in both osteoarthritis and rheumatoid arthritis. The protein possesses a series of domains including N-terminal latent and propetide domains, a C-terminal domain homologous with homopexin, and an internal catalytic domain.

Studies have shown that removal of the N-terminal prosequence of approximately eighty amino acids occurs to convert the proenzyme to the 45 kDa mature enzyme. Furthermore, studies have shown that the C-terminal homopexin homologous domain is not required for proper folding of the catalytic domain or for interaction with an inhibitor. (See, e.g., A. I. Marcy, Biochemistry, 30: 6476-6483 (1991). Thus, the 81-256 amino acid residue internal segment of stromelysin was selected as the protein fragment for use in identifying compounds which bind to and have the potential as acting as inhibitors of stromelysin.

To employ the method of the present invention, it was necessary to prepare the 81-256 fragment (SEQ ID NO:1) of stromelysin in which the peptide backbone was isotopically enriched with and ^{15}N. This was done by inserting a plasmid which coded for the production of the protein fragment into an *E. coli* strain and growing the genetically-modified bacterial strain in a limiting culture medium enriched with $^{15}NH_4Cl$ and ^{13}C-glucose.

The isotopically enriched protein fragment was isolated from the culture medium, purified, and subsequently used as the basis for evaluating the binding of test compounds. The procedures for these processes are described below.

Human skin fibroblasts (ATCC No. CRL 1507) were grown and induced using the procedure described by Clark et al., Archiv. Biochem. and Biophys., 241: 36-45 (1985). Total RNA was isolated from 1 g of cells using a Promega RNAgents® Total RNA Isolation System Kit (Cat.# Z5110, Promega Corp., 2800 Woods Hollow Road, Madison, WI 53711-5399) following the manufacturer's instructions. A 1 μg portion of the RNA was heat-denatured at 80°C for five minutes and then subjected to reverse transcriptase PCR using a GeneAmp® RNA PCR kit (Cat.# N808-0017, Applied Biosystems/Perkin-Elmer, 761 Main Avenue, Norwalk, CT 06859-0156) following the manufacturer's instructions.

Nested PCR was performed using first primers (A) GAAATGAAGAGTC TTCAA (SEQ ID NO:3) and (B) GCGTCCCAGGTTCTGGAG (SEQ ID NO:4) and thirty-five cycles of 94°C, two minutes; 45°C, two minutes; and 72°C three minutes. This was followed by reamplification with internal primers (C) ATACCATGGCCTATCCAT TGGATGGAGC (SEQ ID NO:5) and (D) ATAGGATCCTTAGGTCTCAGGGGA GTCAGG (SEQ ID NO:6) using thirty cycles under the same conditions described immediately above to generate a DNA coding for amino acid residues 1-256 of human stromelysin.

The PCR fragment was then cloned into PCR cloning vector pT7Blue(R) (Novagen, Inc., 597 Science Drive, Madison, WI 53711) according to the manufacturer's instructions. The resulting plasmid was cut with NcoI and BamHI and the stromelysin fragment was subcloned into the Novagen expression vector pET3d (Novagen, Inc., 597 Science Drive, Madison, WI 53711), again using the manufacturer's instructions.

A mature stromelysin expression construct coding for amino acid residues 81-256 plus an initiating methionine was generated from the 1-256 expression construct by PCR amplification. The resulting PCR fragment was first cloned into the Novagen pT7Blue(R) vector and then subcloned into the Novagen pET3d vector, using the manufacturer's instructions in the manner described above, to produce plasmid (pETST-83-256). This final plasmid is identical to that described by Qi-Zhuang et al., Biochemistry, 31: 11231-11235 (1992) with the exception that the present codes for a peptide sequence beginning two amino acids earlier, at position 81 in the sequence of human stromelysin.

Plasmid pETST-83-256 was transformed into *E. coli* strain BL21(DE3)/pLysS (Novagen, Inc., 597 Science Drive, Madison, WI 53711) in accordance with the manufacturer's instructions to generate an expression strain, BL21(DE3)/pLysS/pETST-255-1.

A preculture medium was prepared by dissolving 1.698 g of $Na_2HP_4 \cdot 7H_2O$, 0.45 g of KH_2PO_4, 0.075 g NaCl, 0.150 g $^{15}NH_4Cl$, 0.300 13C-glucose, 300 μL of 1M aqueous $MgSO_4$ solution and 15 μL of aqueous $CaCl_2$ solution in 150 mL of deionized water.

The resulting solution of preculture medium was sterilized and transferred to a sterile 500 mL baffle flask. Immediately prior to inoculation of the preculture medium with the bacterial strain, 150 μL of a solution containing 34 mg/mL of chloramphenicol in 100% ethanol and 1.5 mL of a solution containing 20 mg/mL of ampicillin were added to the flask contents.

The flask contents were then inoculated with 1 mL of glycerol stock of genetically-modified *E. Coli* , strain BL21(DE3)/pLysS/pETST-255-1. The flask contents were shaken (225 rpm) at 37°C until an optical density of 0.65 was observed.

A fermentation nutrient medium was prepared by dissolving 113.28 g of $Na_2HP_4 \cdot 7H_2O$, 30 g of KH_2PO_4, 5 g NaCl and 10 mL of 1% DF-60 antifoam agent in 9604 mL of deionized water. This solution was placed in a New Brunswick Scientific Micros Fermenter (Edison, NJ) and sterilized at 121°C for 40 minutes.

Immediately prior to inoculation of the fermentation medium, the following pre-sterilized components were added to the fermentation vessel contents: 100 mL of a 10% aqueous solution of $^{15}NH_4Cl$, 100 mL of a 10% aqueous solution of ^{13}C-glucose, 20 mL of an aqueous 1M solution of $MgSO_4$, 1 mL of an aqueous 1M $CaCl_2$ solution, 5 mL of an aqueous solution of thiamin hydrochloride (10 mg/mL), 10 mL of a solution containing 34 mg/mL of chloramphenicol in 100% ethanol and 1.9 g of ampicillin dissolved in the chloramphenicol solution. The pH of the resulting solution was adjusted to pH 7.00 by the addition of an aqueous solution of 4N H_2SO_4.

The preculture of *E. Coli* , strain BL21(DE3)/pLysS/pETST-255-1, from the shake-flask scale procedure described above was added to the fermentor contents and cell growth was allowed to proceed until an optical density of 0.48 was achieved. During this process, the fermenter contents were automatically maintained at pH 7.0 by the addition of 4N H_2SO_4 or

4N KOH as needed. The dissolved oxygen content of the fermenter contents was maintained above 55% air saturation through a cascaded loop which increased agitation speed when the dissolved oxygen content dropped below 55%. Air was fed to the fermenter contents at 7 standard liters per minute (SLPM) and the culture temperature was maintained at 37°C throughout the process.

The cells were harvested by centrifugation at 17,000 x g for 10 minutes at 4°C and the resulting cell pellets were collected and stored at -85°C. The wet cell yield was 3.5 g/L. Analysis of the soluble and insoluble fractions of cell lysates by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) revealed that approximately 50% of the ^{15}N-stromelysin was found in the soluble phase.

The isotopically-labeled stromelysin fragment prepared as described above was purified employing a modification of the technique described by Ye et al., Biochemistry, 31: 11231-11235 (1992).

The harvested cells were suspended in 20 mM Tris-HCl buffer (pH 8.0) sodium azide solution containing 1 mM $MgCl_2$, 0.5 mM $ZnCl_2$, 25 units/mL of Benzonase® enzyme, and an inhibitor mixture made up of 4-(2-aminoethyl)-benzenesulfonyl fluoride ("AEBSF"), Leupeptin®, Aprotinin®, and Pepstatin® (all at concentrations of 1 μg/mL. AEBSF, Leupeptin®, Aprotinin®, and Pepstatin® are available from American International Chemical, 17 Strathmore Road, Natick, MA 01760.)

The resulting mixture was gently stirred for one hour and then cooled to 4°C. The cells were then sonically disrupted using a 50% duty cycle. The resulting lysate was centrifuged at 14,000 rpm for 30 minutes and the pellet of insoluble fraction frozen at -80°C for subsequent processing (see below).

Solid ammonium sulfate was added to the supernatant to the point of 20% of saturation and the resulting solution loaded onto a 700 mL phenyl sepharose fast flow ("Q-Sepharose FF") column (Pharmacia Biotech., 800 Centennial Ave., P. O. Box 1327, Piscataway, NJ 08855). Prior to loading, the sepharose column was equilibrated with 50 mM Tris-HCl buffer (pH 7.6 at 4°C), 5 mM $CaCl_2$, and 1 M $(NH_4)_2SO_4$. The loaded column was eluted with a linear gradient of decreasing concentrations of aqueous $(NH_4)_2SO_4$ (from 1 down to 0 M) and increasing concentrations of aqueous $CaCl_2$ (from 5 to 20 mM) in Tris-HCl buffer at pH 7.6.

The active fractions of eluate were collected and concentrated in an Amicon stirred cell (Amicon, Inc., 72 Cherry Hill Drive, Beverly, MA 01915). The concentrated sample was dialyzed overnight in the starting buffer used with the Q-Sepharose FF column, 50 mM Tris-HCl (pH 8.2 at 4°C) with 10 mM $CaCl_2$.

The dialyzed sample was then loaded on the Q-Sepharose FF column and eluted with a linear gradient comprising the starting buffer and 200 mM NaCl. The purified soluble fraction of the isotopically-labeled stromelysin fragment was concentrated and stored at 4°C.

The pellet was solubilized in 8M guanidine-HCl. The solution was centrifuged for 20 minutes at 20,000 rpm and the supernatant was added dropwise to a folding buffer comprising 50 mM Tris-HCl (pH 7.6), 10 mM $CaCl_2$ 0.5 mM $ZnCl_2$ and the inhibitor cocktail of AEBSF, Leupeptin®, Aprotinin®, and Pepstatin® (all at concentrations of 1 μg/mL). The volume of folding buffer was ten times that of the supernatant. The mixture of supernatant and folding buffer was centrifuged at 20,000 rpm for 30 minutes.

The supernatant from this centrifugation was stored at 4°C and the pellet was subjected twice to the steps described above of solubilization in guanidine-HCl, refolding in buffer, and centrifugation. The final supernatants from each of the three centrifugations were combined and solid ammonium sulfate was added to the point of 20% saturation. The resulting solution thus derived from the insoluble fraction was subjected to purification on phenyl Sepharose and Q-Sepharose as described above for the soluble fraction.

The purified soluble and insoluble fractions were combined to produce about 1.8 mg of purified isotopically-labeled stromelysin 81-256 fragment per gram of original cell paste.

B. Human papillomavirus (HPV) E2 Inhibitors

The papillomaviruses are a family of small DNA viruses that cause genital warts and cervical carcinomas. The E2 protein of HPV regulates viral transcription and is required for viral replication. Thus, molecules that block the binding of E2 to DNA may be useful therapeutic agents against HPV. The protein rather than the DNA was chosen as a target, because it is expected that agents with greater selectivity would be found that bind to the protein rather than the DNA.

The DNA-binding domain of human papillomavirus E2 was cloned from the full length DNA that codes for E2 using PCR and overexpressed in bacteria using the T7 expression system. Uniformly ^{15}N-labeled protein was isolated from

bacteria grown on a minimal medium containing ^{15}N-labeled protein was isolated from bacteria grown on a minimal medium containing ^{15}N-labeled ammonium chloride. The protein was purified from the bacterial cell lysate using an S-sepharose FastFlow column pre-equilibrated with buffer (50 mM Tris, 100 mM NaCl, 1 mM EDTA, pH = 8.3).

The protein was eluted with a linear gradient of 100-500 mM NaCl in buffer, pooled, and applied to a Mono-S column at a pH = 7.0. The protein was eluted with a salt gradient (100-500 mM), concentrated to 0.3 mM, and exchanged into a TRIS (50 mM, pH = 7.0 buffered H_2O/D_2O (9/1) solution containing sodium azide (0.5%).

C. RAF

Uniformly ^{15}N-labeled Ras-binding domain of the RAF protein was prepared as described in Emerson et al., Biochemistry, **34** (21): 6911-6918 (1995).

D. FKBP

Uniformly ^{15}N-labeled recombinant human FK binding protein (FKBP) was prepared as described in Logan et al., J. Mol. Biol., **236**: 637-648 (1994).

Example 2

Screening Compounds Using Two-Dimensional $^{15}N/^{1}H$ NMR Correlation Spectral Analysis

The catalytic domain of stromelysin was prepared in accordance with the procedures of Example 1. The protein solutions used in the screening assay contained the uniformly ^{15}N-labeled catalytic domain of stromelysin (0.3 mM), acetohydroxamic acid (500 mM), $CaCl_2$ (20 mM), and sodium azide (0.5%) in a H_2O/D_2O (9/1) TRIS buffered solution (50 mM, pH=7.0).

Two-dimensional $^{15}N/^{1}H$ NMR spectra were generated at 29°C on a Bruker AMX500 NMR spectrometer equipped with a triple resonance probe and Bruker sample changer. The $^{15}N/^{1}H$ HSQC spectra were acquired as 80 x 1024 complex points using sweep widths of 2000 Hz (^{15}N, t^1) and 8333 Hz (^{1}H, t2). A delay of 1 second between scans and 8 scans per free induction decay(fid) were employed in the data collection. All NMR spectra were processed and analyzed on Silicon Graphics computers using in-house-written software.

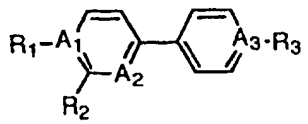
A first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum was acquired for the ^{15}N-labeled stromelysin target molecule as described above. The stromelysin target was then exposed to a database of test compounds. Stock solutions of the compounds were made at 100 mM and 1 M. In addition, a combination library was prepared that contained 8-10 compounds per sample at a concentration of 100 mM for each compound.

The pH of the 1 M stock solution was adjusted with acetic acid and ethanolamine so that no pH change was observed upon a 1/10 dilution with a 100 mM phosphate buffered solution (pH = 7.0). It is important to adjust the pH, because small changes in pH can alter the chemical shifts of the biomolecules and complicate the interpretation of the NMR data.

The compounds in the database were selected on the basis of size (molecular weight = 100-300) and molecular diversity. The molecules in the collection had different shapes (e.g., flat aromatic rings(s), puckered aliphatic rings(s), straight and branched chain aliphatics with single, double, or triple bonds) and diverse functional groups (e.g., carboxylic acids, esters, ethers, amines, aldehydes, ketones, and various heterocyclic rings) for maximizing the possibility of discovering compound that interact with widely diverse binding sites.

The NMR samples were prepared by adding 4 μl of the DMSO stock solution of the compound mixtures that contained each compound at a concentration of 100 mM to 0.4 ml H_2O/D_2O (9/1) buffered solution of the uniformly ^{15}N-labeled protein. The final concentration of each of the compounds in the NMR sample was about 1 mM.

In an initial screen, two compounds were found that bind to the catalytic domain of stromelysin. Both of these compounds contain a biaryl moiety. Based on these initial hits, structurally similar compounds were tested against stromelysin. The structure of those biaryl compounds is represented by the structure I, below. (See Table 1 for definitions of R_1-R_3 and A_1-A_3).



In the second round of screening, binding was assayed both in the absence and in the presence of saturating amounts of acetohydroxamic acid (500 mM).

Many of the biaryl compounds were found to bind the catalytic domain of stromelysin. FIG. 4 shows a representative two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum before and after exposure of stromelysin to a biaryl test compound. It can be seen from FIG. 4 that the compound caused chemical shifts of ^{15}N-sites such as those designated W124, T187, A199 and G204.

These sites correspond to a tryptophan (Trp) residue at position 124, a threonine (Thr) at position 187, an alanine (Ala) at position199, and a glycine (Gly) at position 204 of SEQ ID NO. 1. FIG. 9 shows the correlation between the NMR binding data and a view of the NMR-derived three-dimensional structure of the catalytic domain of stromelysin. The ability to locate the specific binding site of a particular ligand is an advantage of the present invention.

Some compounds only bound to stromelysin in the presence of hydroxamic acid. Thus, the binding affinity of some compounds was enhanced in the presence of the hydroxamic acid (i. e. cooperative). These results exemplify another important capability of the present screening assay: the ability to identify compounds that bind to the protein in the presence of other molecules.

Various biaryl compounds of structure I were tested for binding to stromelysin at differing concentrations. The $^{15}N/^{1}H$ spectra generated at each concentration were evaluated to quantify differences in the spectra as a function of compound concentration. A binding or dissociation constant (K_D)was calculated, using standard procedures well known in the art, from those differences. The results of this study are shown in Table 1. The values for R1-R3 and A1-A3 in Table 1 refer to the corresponding positions in the structure I, above.

Table 1

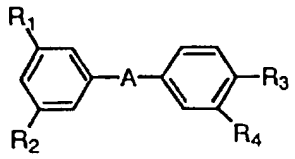
Compound No.	R_1	R_2	R_3	A_1	A_2	A_3	K_D(mM)
1	H	OH	H	C	C	C	1.1
2	CH_2OH	H	H	C	C	C	3.2
3	Br	H	OH	C	C	C	1.3
4	H	H	H	N	N	C	1.6
5	CHO	H	H	C	C	C	1.7
6	OCH_3	NH_2	H	C	C	C	0.4
7	H	H	H	N	C	C	0.2

8	$OCOCH_3$	H	H	C	C	C	0.3
9	OH	H	OH	C	C	C	0.16
10	H	H	H	N	C	N	0.4
11	OH	H	H	C	C	C	0.3
12	OH	H	CN	C	C	C	0.02

The data in Table 1 show the utility of a process of the present invention in determining dissociation or binding constants between a ligand and a target molecule.

Another advantage of an NMR screening assay of the present invention is the ability to correlate observed chemical shifts from the two-dimensional $^{15}N/^{1}H$ NMR correlation spectra with other spectra or projections of target molecule configuration. The results of a representative such correlation are shown in FIG. 9, which depicts regions within the polypeptide at which binding with the substrate molecule is most likely occurring. In this Figure, the apparent binding regions in stromelysin are shown for Compound 1 (from Table 1).

Compounds from the database were screened in a similar manner for binding to the DNA-binding domain of the E2 protein. Those compounds had the structure II below, where R_1-R_4 and A are defined in Table 2.



NMR experiments were performed at 29°C on a Bruker AMX500 NMR spectrometer equipped with a triple resonance probe and Bruker sample changer. The ^{15}N-/^{1}H HSQC spectra were acquired as 80 x 1024 complex points using sweep widths of 2000 Hz (^{15}N,t_1) and 8333 Hz (^{1}H, t_2). A delay of 1 second between scans and 4 scans per free induction decay were employed in the data collection. All NMR spectra were processed and analyzed on Silicon Graphics computers.

FIGs. 2 and 3 show representative two-dimensional $^{15}N/^{1}H$ NMR correlation spectra before and after exposure of the DNA-binding domain of E2 to a first and second test compound, respectively.

It can be seen from FIG. 2 that the first test compound caused chemical shifts of ^{15}N-sites such as those designated I15, Y21, R22 and L23. Those sites correspond to an isoleucine (Ile) residue at position 15, a tyrosine residue (Tyr) at position 21, an arginine (Arg residue at position 22 and a leucine (Leu) residue at position 23 of SEQ ID NO. 6.

It can be seen from FIG. 3 that the second test compound caused chemical shifts in the particular ^{15}N-sites designated I6, G11, H38, and T52. Those sites correspond to an isoleucine (Ile) residue at position 6, a glycine (Gly) residue at position 11, a histidine (His) residue at position 38 and a threonine (Thr) at position 52 of SEQ ID NO. 6.

FIGs. 7 and 8 show the correlation between those NMR binding data and a view of the NMR-derived three-dimensional structure of the DNA-binding domain of E2.

Several structurally similar compounds caused chemical shift changes of the protein signals when screened at a concentration of 1 mM. Two distinct sets of amide resonances were found to change upon the addition of the compounds: one set of signals corresponding to amides located in the ß-barrel formed between the two monomers and a second set corresponding to amides located near the DNA-binding site.

For example, compounds containing two phenyl rings with a carboxylic acid attached to the carbon linking the two rings only caused chemical shift changes to the amides in the DNA-binding site. In contrast, benzophenones and phenoxyphenyl-containing compounds only bound to the ß-barrel. Other compounds caused chemical shift changes of both sets of signals but shifted the signals in each set by different amounts, suggesting the presence of two distinct binding sites.

By monitoring the chemical shift changes as a function of ligand concentration, binding constants for the two binding sites were also measured. The results of those studies are summarized below in Table 2.

Table 2

Comp. No.	A	R_1	R_2	R_3	R_4	DNA K_D(mM)	β-barrel K_D(mM)	Filter binding assay
13	CO	H	H	H	OH	>50	0.6	-
14	O	H	H	H	CH_2OH	>50	2.0	-
15	-[a]	H	H	COO	H	2.0	>50	+
16	-[a]	Cl	Cl	COO	H	0.1	>50	+
17	-[a]	H	H	CH_2COO	H	4.2	4.9	+
18	-[a]	H	H	CH=CHCOO	H	1.2	6.2	+
19	O	H	H	$CH_2CH_2CH(CH_3)$ $-CH_2COO$	H	0.5	0.2	+
20	O	H	H	$COCH_2CH_2COO$	H	2.7	4.8	+

[a] a dash (-) for A indicates no atom (i.e. byphenyl linkage)

Uniformly ^{15}N-labeled Ras-binding domain of the RAF protein was prepared as described in Example 1 and screened using two-dimensional $^{15}N/^{1}H$ NMR correlation spectral analysis in accordance with the NMR procedures described above. The results of a representative study are shown in FIG. 5, which depicts two-dimensional $^{15}N/^{1}H$ NMR correlation spectra both before and after exposure to a test compound.

Uniformly ^{15}N-labeled FKBP was prepared as described in Example 1 and screened using two-dimensional $^{15}N/^{1}H$ NMR correlation spectral analysis in accordance with the NMR procedures described above. The results of a representative study are shown in FIG. 6, which depicts two-dimensional $^{15}N/^{1}H$ NMR correlation spectra both before and after exposure to a test compound.

Example 3

Comparison of NMR, Enzymatic, Filter Binding and Gel Shift Screening Assays

Studies were performed to compare binding constants of ligands to various biomolecules, determined by the NMR method of the present invention, to similar results obtained from prior art methods.

In a first study, binding constants were determined, both by the NMR method of the present invention, and by a prior art enzymatic assay. The target molecule was the catalytic domain of stromelysin prepared in accordance with the

procedures of Example 1. The NMR binding constants, K_D, were derived using two-dimensional $^{15}N/^{1}H$ NMR correlation spectroscopy as described in Example 2. The K_D values so obtained were compared to an inhibition constant K_I as determined in an enzymatic assay.

The enzymatic assay measured the rate of cleavage of a fluorogenic substrate by following the fluorescence increase upon peptide cleavage which causes a separation between the fluorophore and quencher. Enzymatic activity was measured using a matrix of different concentrations of acetohydroxamic acid and biaryl compounds. The assay is a modification of the method described by H. Weingarten, *et al.* in Anal. Biochem., **147**: 437-440 (1985) employing the fluorogenic substrate properties described by E. Matayoshi, *et al.* in Science: **247**: 954-958 (1990).

Eight acetohydroxamic acid concentrations were used ranging from 0.0 to 1.0 M, and six compound concentrations were used, resulting in a total of 48 points. Individual compound concentration varied due to solubility and potency.

All NMR measurements were performed in the presence of 500 mM acetohydroxamic acid, except for the titration of acetohydroxamic acid itself. Dissociation constants were obtained from the dependence of the observed chemical shift changes upon added ligand. Inhibition constants were then obtained from the inhibition data using standard procedures.

The results of these studies are summarized below in Table 3, which shows the comparison of NMR-derived dissociation constants (K_D) with inhibition constants measured in the enzyme assay (K_I) using a fluorogenic substrate.

Table 3

Compound No.	NMR K_D (mM)	Assay K_I (mM)
4	1.6	7.4
7	0.17	0.32
9	0.16	0.70
10	0.40	1.8
12	0.02	0.11
Acetohydroxamic acid	17.0	21.1

The data in Table 3 show that a NMR process of the present invention provides a rapid, efficient and accurate way of determining dissociation or binding constants of ligands to target biomolecules. Comparison of the binding constants determined by the two methods result in the same ranking of potencies of the compounds tested. That is, while the values for a given substrate as determined by the two methods are not equal, they are proportional to one another.

In a second study, the results for binding of the DNA-binding domain of E2 to its target DNA were obtained by prior art methods and compared with results obtained by the method of the present invention. The target was the DNA-binding domain of E2, prepared in accordance with the procedures of Example 1. NMR screening assays and NMR processes for determining ligand dissociation constants were performed as set forth above in Example 2.

The binding constant from the NMR process was compared to the results of a physical, filter binding assay that measured binding of DNA to the target. The high-throughput filter binding assay was performed using E2, prepared according to Example 2 above. The ^{33}P-labeled DNA construct comprised a 10,329 base pair plasmid formed by inserting the HPV-11 genome, containing three high affinity and one low affinity E2 binding sites, into the PSP-65 plasmid (Promega, Madison, WI).

The binding affinities at the different sites as determined by NMR were compared for a subset of the compounds to the inhibition of E2 binding to DNA as measured in the filter binding assay. As shown in Table 2 above, the activities determined in the filter binding assay correlated closely with the binding affinities calculated from the amides of the DNA-binding site but not to the affinities measured for the ß-barrel site. This is consistent with the relative locations of each site.

In an alternative study, a comparison of the NMR-determined binding results was made with similar results obtained by a prior art gel-shift assay using techniques well known in the art. The gel-shift assay was performed using a GST fusion protein which contained full length E2 and a ^{33}P-labeled 62 base pair DNA fragment containing two E2 binding sites.

The method identified numerous compounds which gave positive results in the gel-shift assay. Some of these positive results, however, were believed to be due to binding to the DNA, since in these cases, no binding to the E2 protein was observed using the NMR method of this invention. These compounds were shown to indeed bind to DNA rather than to E2, as evidenced by changes in the

chemical shifts of the DNA rather than the protein upon the addition of the compounds. These data show that yet another advantage of the present invention is the ability to minimize the occurrence of false positives.

Example 4

Design of a potent, non-peptide inhibitor of stromelysin

Studies were performed to design new ligands that bound to the catalytic domain of stromelysin. Because stromelysin undergoes autolysis, an inhibitor was sought to block the degradation of stromelysin. That inhibitor would facilitate the screening of other potential ligands that bind to other sites on the enzyme.

The criteria used in selecting compounds in the screening for other binding sites was based primarily on the size of the ligand. The smallest ligand was sought that had enough solubility to saturate (>98% occupancy of enzyme) and inhibit the enzyme.

The cloning, expression, and purification of the catalytic domain of stromelysin was accomplished using the procedures set forth in Example 1. An initial step in the design of the new ligand was the identification of a first ligand that bound to the stromelysin target. Such identification was carried out in accordance with a two-dimensional $^{15}N/^{1}H$ NMR correlation screening process as disclosed above.

A variety of hydroxamic acids of the general formula R-(CO)NHOH were screened for binding to stromelysin using the procedures set forth in Example 2. Of the compounds tested, acetohydroxamic acid [CH3(CO)NHOH] best satisfied the selection criteria: it had a binding affinity for stromelysin of 17 mM and had good water solubility. At a concentration of 500 mM, acetohydroxamic acid inhibited the degradation of the enzyme, allowing the screening of other potential ligands.

The second step in the design process was the identification of a second ligand that bound to the target stromelysin at a site different from the binding site of acetohydroxamic acid. This was accomplished by screening compounds for their ability to bind stromelysin in the presence of saturating amounts of acetohydroxamic acid. Details of procedures and results of this second identification step are set forth above in Example 2.

The compound identified as a second ligand from these studies and used in subsequent design steps was the compound designated as Compound #4 in Table 1 (See Example 2).

The next step in the design process was to construct a ternary complex of the target stromelysin, the first ligand and the second ligand. This was accomplished by exposing the stromelysin target to the two ligands under conditions that resulted in complex formation. The three-dimensional structure of the ternary complex was then determined using NMR spectroscopy as described below.

The ^{1}H, ^{13}C, and ^{15}N backbone resonances of stromelysin in the ternary complex were assigned from an analysis of several 3D double- and triple-resonance NMR spectra (A. Bax. *et al.*. Acc. Chem. Res., **26**: 131-138 (1993)). The C^{α} resonances of adjacent spin systems were identified from an analysis of three-dimensional (3D) HNCA (L. Kay *et al.*, J. Magn. Reson., **89**: 496-514 (1990)) and HN(CO)CA (A. Bax, *et al.*, J. Bio. NMR, **1**: 99 (1991)) spectra recorded with identical spectral widths of 1773 Hz (35.0 ppm), 3788 Hz (30.1 ppm), and 8333 Hz (16.67 ppm) in the $F_1(^{15}N)$, $F_2(^{13}C)$ and $F_3(^{1}H)$ dimensions, respectively.

The data matrix was $38(t_1) \times 48(t_2) \times 1024(t_3)$ complex points for the HNCA spectrum, and $32(t_1) \times 40(t_2) \times 1024(t_3)$ complex points for the HN(CO)CA spectrum. Both spectra were acquired with 16 scans per increment. A 3D CBCA(CO)NH spectrum (S. Grzesiek, *et al.*, J. Am. Chem. Soc., **114**: 6261-6293 (1992)) was collected with $32(t_1, {}^{15}N) \times 48(t_2, {}^{13}C) \times 1024(t_3, {}^{1}H)$ complex points and 32 scans per increment. Spectral widths were 1773 Hz (35.0 ppm), 7575.8 Hz (60.2 ppm), and 8333 Hz (16.67 ppm) in the ^{15}N, ^{13}C and ^{1}H dimensions, respectively.

For all three spectra, the ^{1}H carrier frequency was set on the water resonance and the ^{15}N carrier frequency was at 119.1 ppm. The ^{13}C carrier frequency was set to 55.0 ppm in HNCA and HN(CO)CA experiments, and 46.0 ppm in the CBCA(CO)NH experiment.

The backbone assignments were confirmed from an analysis of the crosspeaks observed in an ^{15}N-separated 3D NOESY-HSQC spectrum and a 3D HNHA-J spectrum. The ^{15}N-separated 3D NOESY-HSQC spectrum (S. Fesik, *et al.*, J. Magn. Reson., **87**: 588-593 (1988)); D. Marion, *et al.*, J. Am. Chem. Soc., **111**: 1515-1517 (1989)) was collected with a mixing time of 80 ms. A total of $68(t_1, {}^{15}N) \times 96(t_2, {}^{1}H) \times 1024(t_3, {}^{1}H)$ complex points with 16 scans per increment were collected, and the spectral widths were 1773 Hz (35.0 ppm) for the ^{15}N dimension, 6666.6 Hz (t_2, ^{1}H, 13.3 ppm), and 8333 Hz (16.7 ppm) for the ^{1}H dimension.

The 3D HNHA-J spectrum (G. Vuister, *et al.*, J. Am. Chem. Soc., **115**: 7772-7777 (1993)), which was also used to obtain $^{3}J_{HNH\alpha}$ coupling constants, was acquired with 35(t_1, ^{15}N) x 64(t_2, ^{1}H) x 1024(t_3, ^{1}H) complex points and 32 scans per increment. Spectral widths and carrier frequencies were identical to those of the ^{15}N-separated NOESY-HSQC spectrum. Several of the H^{β} signals were assigned using the HNHB experiment. The sweep widths were the same as in the ^{15}N-separated NOESY-HSQC spectrum that was acquired with 32(t_1, ^{15}N) x 96(t_2, ^{1}H) x 1024 (t_3, ^{1}H) complex points.

The ^{1}H and ^{13}C chemical shifts were assigned for nearly all sidechain resonances. A 3D HCCH-TOCSY spectrum (L. Kay, *et al.*, J. Magn. Reson., **101b**: 333-337 (1993)) was acquired with a mixing time of 13 ms using the DIPSI-2 sequence (S. Rucker, *et al.*, Mol. Phys., **68**: 509 (1989)) for ^{13}C isotropic mixing. A total of 96 (t_1, ^{13}C) x 96(t_2, ^{1}H) x 1024(t_3, ^{1}H) complex data points were collected with 16 scans per increment using a spectral width of 10638 Hz (70.8 ppm, w_1), 4000 Hz (6.67 ppm, w_2), and 4844 (8.07 ppm, w_3). Carrier positions were 40 ppm, 2.5 ppm, and at the water frequency for the ^{13}C, indirectly detected ^{1}H, and observed ^{1}H dimensions, respectively.

Another 3D HCCH-TOCSY study was performed with the ^{13}C carrier at 122.5 ppm to assign the aromatic residues. The spectra were collected with 36(t_1, ^{13}C) x 48(t_2, ^{1}H) x 1024 (t_3, ^{1}H) complex points with spectral widths of 5263 Hz (35.0 ppm, w_1), 3180 Hz (5.30 ppm, w_2), and 10,000 (16.7 ppm, w_3). Carrier positions were 122.5 ppm, 7.5 ppm, and at the water frequency for the ^{13}C, indirectly detected ^{1}H, and observed ^{1}H dimensions, respectively.

A ^{13}C-separated 3D NOESY-HMQC spectrum (S. Fesik, *et al.*, J. Magn. Reson., **87**: 588-593 (1988)); D. Marion, *et al.*, J. Am. Chem. Soc., **111**: 1515-1517 (1989)) was recorded using a mixing time of 75 ms. A total of 80 (t_1, ^{13}C) x 72 (t_2, ^{1}H) x 1024 (t_3, ^{1}H) complex data points with 16 scans per increment were collected over spectral widths of 10638 Hz (70.49 ppm, w_1), 6666.6 Hz (13.3 ppm, w_2), and 8333.3 Hz (16.67 ppm, w_3). The ^{1}H carrier frequencies were set to the water resonance, and the ^{13}C carrier frequency was placed at 40.0 ppm.

Stereospecific assignments of methyl groups of the valine and leucine residues were obtained by using a biosynthetic approach (Neri *et al.*, Biochem., **28**: 7510-7516 (1989)) on the basis of the ^{13}C-^{13}C one-bond coupling pattern observed in a high-resolution ^{1}H, ^{13}C-HSQC spectrum (G. Bodenhausen, *et al.*, J. Chem. Phys. Lett., **69**: 185-189 (1980)) of a fractionally ^{13}C-labeled protein

sample. The spectrum was acquired with 200(^{13}C, t_1) x 2048(^{1}H, t_2) complex points over spectral widths of 5000 Hz (39.8 ppm, ^{13}C) and 8333 Hz (16.7 ppm, ^{1}H). Carrier positions were 20.0 ppm for the ^{13}C dimension, and at the water frequency for the ^{1}H dimension.

To detect NOEs between the two ligands and the protein, a 3D ^{12}C-filtered, ^{13}C-edited NOESY spectrum was collected. The pulse scheme consisted of a double ^{13}C-filter sequence (A. Gemmeker, *et al.*, J. Magn. Reson., **96**: 199-204 (1992)) concatenated with a NOESY-HMQC sequence (S. Fesik, *et al.*, J. Magn. Reson., **87**: 588-593 (1988)); D. Marion, *et al.*, J. Am. Chem. Soc., **111**: 1515-1517 (1989)) . The spectrum was recorded with a mixing time of 80 ms, and a total of 80 (t_1, ^{13}C) x 80 (t_2, ^{1}H) x 1024 (t_3, ^{1}H) complex points with 16 scans per increment. Spectral widths were 8865 Hz (17.73 ppm, w_1), 6667 Hz (13.33 ppm, w_2), and 8333 Hz (16.67 ppm, w_3), and the carrier positions were 40.0 ppm for the carbon dimension and at the water frequency for both proton dimensions.

To identify amide groups that exchanged slowly with the solvent, a series of ^{1}H, ^{15}N-HSQC spectra (G. Bodenhausen, *et al.*, J. Chem. Phys. Lett., **69**: 185-189 (1980)) were recorded at 25°C at 2 hr intervals after the protein was exchanged into D_2O. The acquisition of the first HSQC spectrum was started 2 hrs. after the addition of D_2O.

All NMR spectra were recorded at 25°C on a Bruker AMX500 or AMX600 NMR spectrometer. The NMR data were processed and analyzed on Silicon Graphics computers. In all NMR experiments, pulsed field gradients were applied where appropriate as described (A. Bax, *et al.*, J. Magn. Reson., **99**: 638 (1992)) to afford the suppression of the solvent signal and spectral artifacts. Quadrature detection in indirectly detected dimensions was accomplished by using the States-TPPI method (D. Marion, *et al.*, J. Am. Chem. Soc., **111**: 1515-1517 (1989)). Linear prediction was employed as described (E. Olejniczak, *et al.*, J. Magn. Reson., **87**: 628-632 (1990)).

The derived three-dimensional structure of the ternary complex was then used to define the spatial orientation of the first and second ligands to each other as well as to the target stromelysin molecule.

Distance restraints derived from the NOE data were classified into six categories based on the NOE cross peak intensity and given a lower bound of 1.8 Å and upper bounds of 2.5 Å, 3.0 Å, 3.5 Å, 4.0 Å, 4.5 Å, and 5.0 Å, respectively. Restraints for ϕ torsional angles were derived from $^{3}J_{HNH\alpha}$

coupling constants measured from the 3D HNHA-J spectrum (G. Vuister, *et al.*, J. Am. Chem. Soc., **115**: 7772-7777 (1993)). The ϕ angle was restrained to 120%±40% for $^{3}J_{HNH\alpha} > 8.5$ Hz, and 60%±40% for $^{3}J_{HNH\alpha} < 5$ Hz.

Hydrogen bonds, identified for slowly exchanging amides based on initial structures, were defined by two restraints: 1.8-2.5 Å for the H-O distance and 1.8-3.3 Å for the N-O distance. Structures were calculated with the X-PLOR 3.1 program (A. Brünger, "XPLOR 3.1 Manual," Yale University Press, New Haven, 1992) on Silicon Graphics computers using a hybrid distance geometry-simulated annealing approach (M. Nilges, *et al.*, FEBS Lett., **229**: 317-324 (1988)).

A total of 1032 approximate interproton distance restraints were derived from the NOE data. In addition, 21 unambiguous intermolecular distance restraints were derived from a 3D 12C-filtered, 13C-edited NOESY spectrum. Of the 1032 NOE restraints involving the protein, 341 were intra-residue, 410 were sequential or short-range between residues separated in the primary sequence by less than five amino acids, and 281 were long-range involving residues separated by at least five residues.

In addition to the NOE distance restraints, 14 ϕ dihedral angle restraints were included in the structure calculations that were derived from three-bond coupling constants ($^{3}J_{HNH\alpha}$) determined from an HNHA-J spectrum (G. Vuister, *et al.*, J. Am. Chem. Soc., **115**: 7772-7777 (1993)). The experimental restraints also included 120 distance restraints corresponding to 60 hydrogen bonds. The amides involved in hydrogen bonds were identified based on their characteristically slow exchange rate, and the hydrogen bond partners from initial NMR structures calculated without the hydrogen bond restraints. The total number of non-redundant, experimentally-derived restraints was 1166.

The structures were in excellent agreement with the NMR experimental restraints. There were no distance violations greater than 0.4 Å, and no dihedral angle violations greater than 5 degrees. In addition, the simulated energy for the van der Waals repulsion term was small, indicating that the structures were devoid of bad inter-atomic contacts.

The NMR structures also exhibited good covalent bond geometry, as indicated by small bond-length and bond-angle deviations from the corresponding idealized parameters. The average atomic root mean square deviation of the 8 structures for residues 93-247 from the mean coordinates was 0.93 Å for backbone atoms (C^{a}, N, and C'), and 1.43 Å for all non-hydrogen atoms.

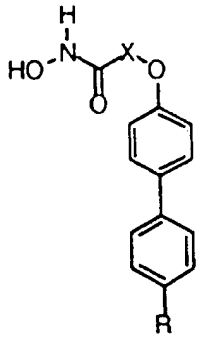
A ribbon plot of the ternary complex involving stromelysin, acetohydroxamic acid (the first ligand), and the second ligand is shown in Fig 10. The structure is very similar to the global fold of other matrix metalloproteinases and consists of a five-stranded ß-sheet and three a-helices.

The catalytic zinc was located in the binding pocket. It was coordinated to three histidines and the two oxygen atom of acetohydroxamic acid. A biaryl group of the second ligand was located in the S1' pocket between the second helix and the loop formed from residues 218-223. This deep and narrow pocket is lined with hydrophobic residues which make favorable contacts with the ligand.

Based on the three-dimensional structure of the ternary complex as determined above and the structure/activity relationships observed for the binding to stromelysin of structural analogs of the second ligand (i.e., other biaryl compounds), new molecules were designed that linked together the acetohydroxamic acid to biaryls.

As shown in Table 4 below, the initial biaryls chosen contained an oxygen linker and the absence or presence of CN para to the biaryl linkage. Initial linkers contained varying lengths of methylene units. Means for linking compounds with linkers having varying lengths of methylene units are well known in the art.

Table 4



Compound	X	R	Stromelysin Inhibition
21	$(CH_2)_2$	H	0.31 μM
22	$(CH_2)_3$	H	110 μM
23	$(CH_2)_4$	H	38%@100 μM

24	$(CH_2)_5$	H	43%@100 μM
25	$(CH_2)_2$	CN	0.025 μM
26	$(CH_2)_3$	CN	3.4 μM
27	$(CH_2)_4$	CN	3.5 μM
28	$(CH_2)_5$	CN	1.7 μM

As expected based on the better binding of the CN substituted biaryls to stromelysin, the CN derivatives exhibited better stromelysin inhibition. The compound that exhibited the best inhibition of stromelysin contained a linker with two methylene units.

The present invention has been described with reference to preferred embodiments. Those embodiments are not limiting of the claims and specification in any way. One of ordinary skill in the art can readily envision changes, modifications and alterations to those embodiments that do not depart from the scope and spirit of the present invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Fesik, Stephen W.
Hajduk, Philip J.

(ii) TITLE OF INVENTION: Use of Nuclear Magnetic Resonance to
Identifyify Ligands to Target Biomolecules

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Steven F. Weinstock, Dept. 377 AP6D, Abbott Laboratories
(B) STREET: 100 Abbott Park Road
(C) CITY: Abbott Park
(D) STATE: Illinois
(E) COUNTRY: USA
(F) ZIP: 60064-3500

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Janssen, Jerry F.
(B) REGISTRATION NUMBER: 29,175

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (708) 937-4558
(B) TELEFAX: (708) 938-7742

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 174 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Phe Arg Thr Phe Pro Gly Ile Pro Lys Trp Arg Lys Thr
1 5 10

His Leu Thr Tyr Arg Ile Val Asn Tyr Thr Pro Asp Leu
15 20 25

Pro Lys Asp Ala Val Asp Ser Ala Val Glu Lys Ala Leu
30 35

Lys Val Trp Glu Glu Val Thr Pro Leu Thr Phe Ser Arg
40 45 50

Leu Tyr Glu Gly Glu Ala Asp Ile Met Ile Ser Phe
55 60

Ala Val Arg Glu His Gly Asp Phe Tyr Pro Phe Asp Gly
65 70 75

Pro Gly Asn Val Leu Ala His Ala Tyr Ala Pro Gly Pro
80 85 90

Gly Ile Asn Gly Asp Ala His Phe Asp Asp Asp Glu Gln
95 100

Trp Thr Lys Asp Thr Thr Gly Thr Asn Leu Phe Leu Val
105 110 115

Ala Ala His Glu Ile Gly His Ser Leu Gly Leu Phe
120 125

His Ser Ala Asn Thr Glu Ala Leu Met Tyr Pro Leu Tyr
130 135 140

His Ser Leu Thr Asp Leu Thr Arg Phe Arg Leu Ser Gln
145 150

Asp Asp Ile Asn Gly Ile Gln Ser Leu Tyr Gly Pro Pro
155 160 165

Pro Asp Ser Pro Glu Thr Pro
170

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 83 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Thr Thr Pro Ile Ile His Leu Lys Gly Asp Ala
5 10

Asn Ile Leu Leu Cys Leu Arg Tyr Arg Leu Ser Lys Tyr
15 20 25

Lys Gln Leu Tyr Glu Gln Val Ser Ser Thr Trp His Trp
30 35

Thr Cys Thr Asp Gly Lys His Lys Asn Ala Ile Val Thr
40 45 50

Leu Thr Tyr Ile Ser Thr Ser Gln Arg Asp Asp Phe Leu
55 60 65

Asn Thr Val Lys Ile Pro Asn Thr Val Ser Val Ser Thr
70 75

Gly Tyr Met Thr Ile
80

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAAATGAAGA GTCTTCAA 18

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGTCCCAGG TTCTGGAG 18

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATACCATGGC CTATCCATTG GATGGAGC 28

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATAGGATCCT TAGGTCTCAG GGGAGTCAGG 30

WHAT IS CLAIMED IS:

1. A process of screening compounds to identify compounds that are ligands that bind to a specific target molecule comprising the steps of:

a) generating a first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of a ^{15}N-labeled target molecule;

b) exposing the labeled target molecule to one or a mixture of chemical compounds;

c) generating a second two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of the labeled target molecule that has been exposed to one or a mixture of compounds in step (b); and

d) comparing said first and second two-dimensional $^{15}N/^{1}H$ NMR correlation spectra to determine differences between said first and said second spectra, the differences identifying the presence of one or more compounds that are ligands which have bound to the target molecule.

2. The process of claim 1 wherein the ^{15}N-labeled target molecule is exposed to a mixture of chemical compounds in step (b), further comprising the steps subsequent to step d) of

e) exposing the ^{15}N-labeled target molecule individually to each compound of said mixture,

f) generating a two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of the labeled target molecule that has been individually exposed to each compound; and

g) comparing each spectrum generated in step f) to said first spectrum to determine differences in any of those compared spectra, the differences identifying the presence of a compound that is a ligand which has bound to the target molecule.

3. The process of claim 1 wherein the differences in the two-dimensional $^{15}N/^{1}H$ NMR correlation spectra are chemical shifts at particular ^{15}N-labeled sites in the target molecule and chemical shifts in protons attached to those ^{15}N-labeled sites.

4. The process of claim 1 wherein the target molecule is a polypeptide.

5. A process of determining the dissociation constant between a target molecule and a ligand that binds to that target molecule comprising the steps of:

a) generating a first two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum of a ^{15}N-labeled target molecule;

b) exposing the labeled target molecule to various concentrations of a ligand;

c) generating a two-dimensional $^{15}N/^{1}H$ NMR correlation spectrum at each concentration of ligand from step (b);

d) comparing each spectrum from step (c) both to the first spectrum from step (a) and to all other spectra from step (c) to quantify differences in those spectra as a function of changes in ligand concentration; and

e) calculating the dissociation constant between the target molecule and the ligand from those differences according to the equation:

$$K_D = \frac{([P]_o - x)([L]_o - x)}{x}$$

where $[P]_O$ is the total molar concentration of target molecule;

$[L]_O$ is the total molar concentration of ligand; and

x is the molar concentration of the bound species determined according to the equation:

$$x = \frac{\delta_{obs} - \delta_{free}}{\Delta}$$

where δ_{obs} and δ_{free} are the chemical shift values for the target molecule determined at each concentration of ligand and for the target molecule in the absence of ligand, respectively, and Δ is the difference between the chemical shift at saturating amounts of ligand and δ_{free}.

6. The process of claim 5 wherein the target molecule is a polypeptide.

7. The process of claim 5 further comprising the step of binding the labeled target molecule to a second ligand before step (a).

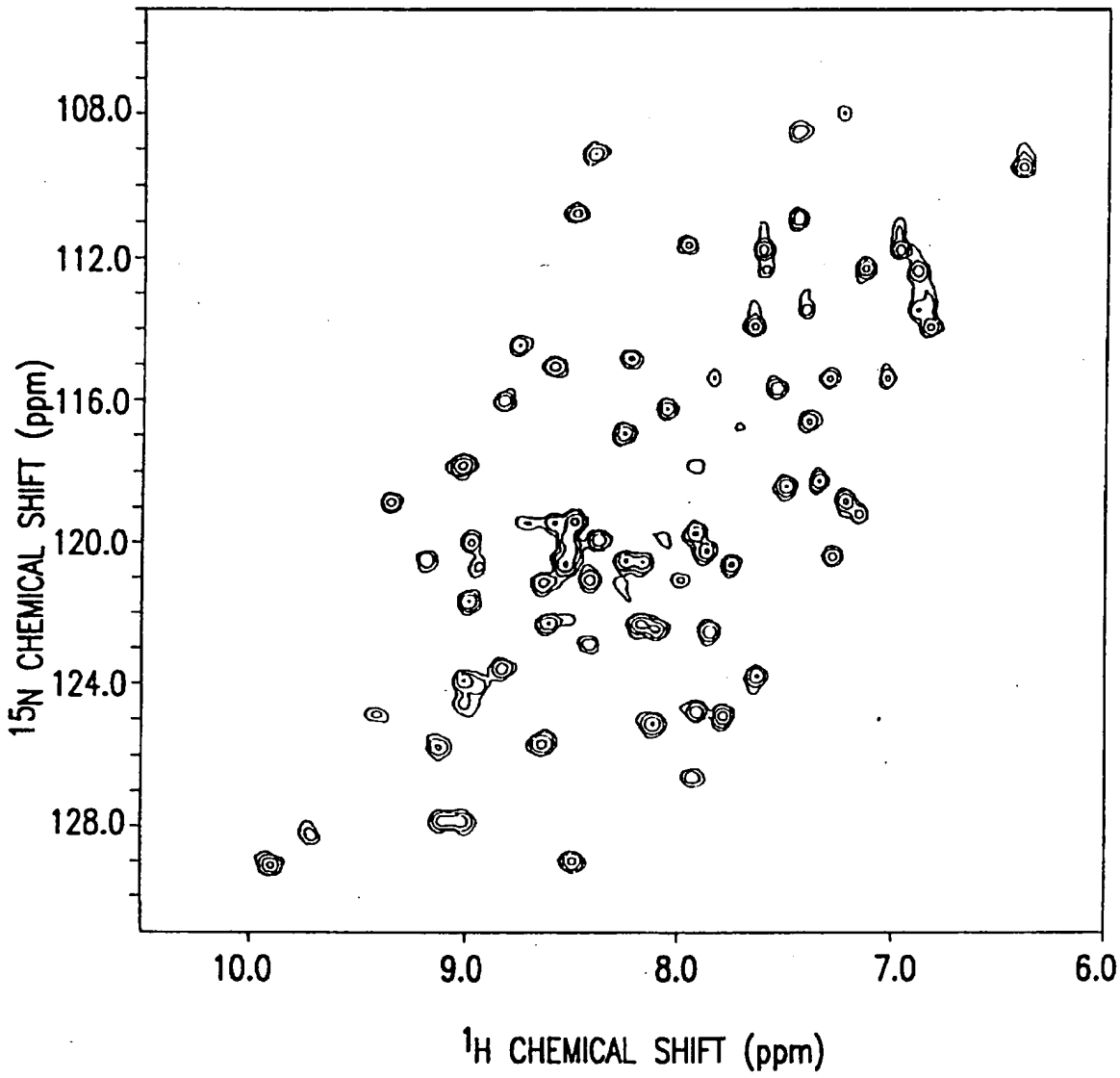

108.0
112.0
116.0
120.0
124.0
128.0
15N CHEMICAL SHIFT (ppm)
10.0
9.0
8.0
7.0
6.0
1H CHEMICAL SHIFT (ppm)

FIG.1

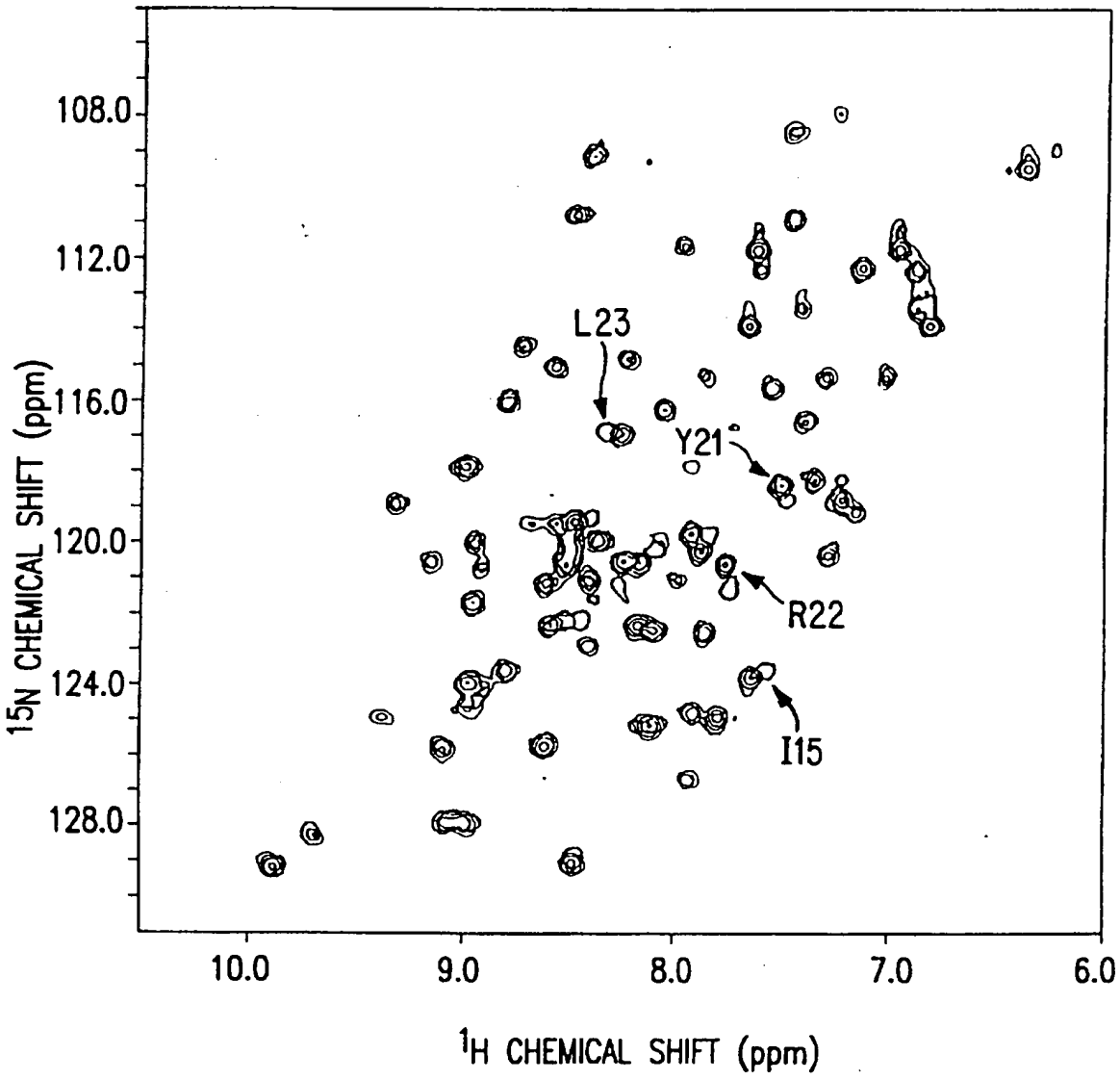

108.0
112.0
116.0
120.0
124.0
128.0
15N CHEMICAL SHIFT (ppm)
L23
Y21
R22
I15
10.0
9.0
8.0
7.0
6.0
1H CHEMICAL SHIFT (ppm)

FIG.2

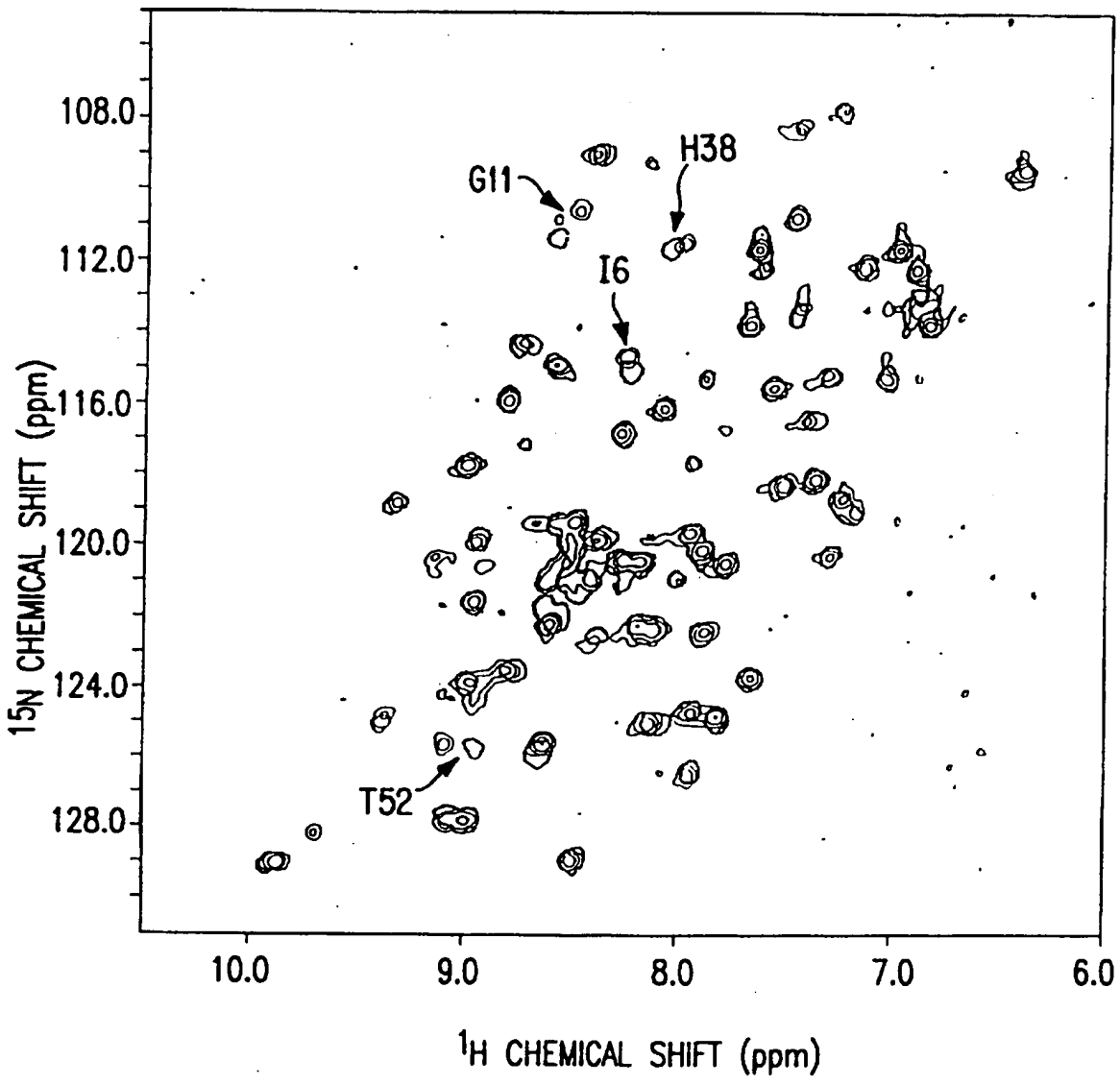

108.0
112.0
116.0
120.0
124.0
128.0
15N CHEMICAL SHIFT (ppm)
G11
H38
I6
T52
10.0
9.0
8.0
7.0
6.0
1H CHEMICAL SHIFT (ppm)

FIG.3

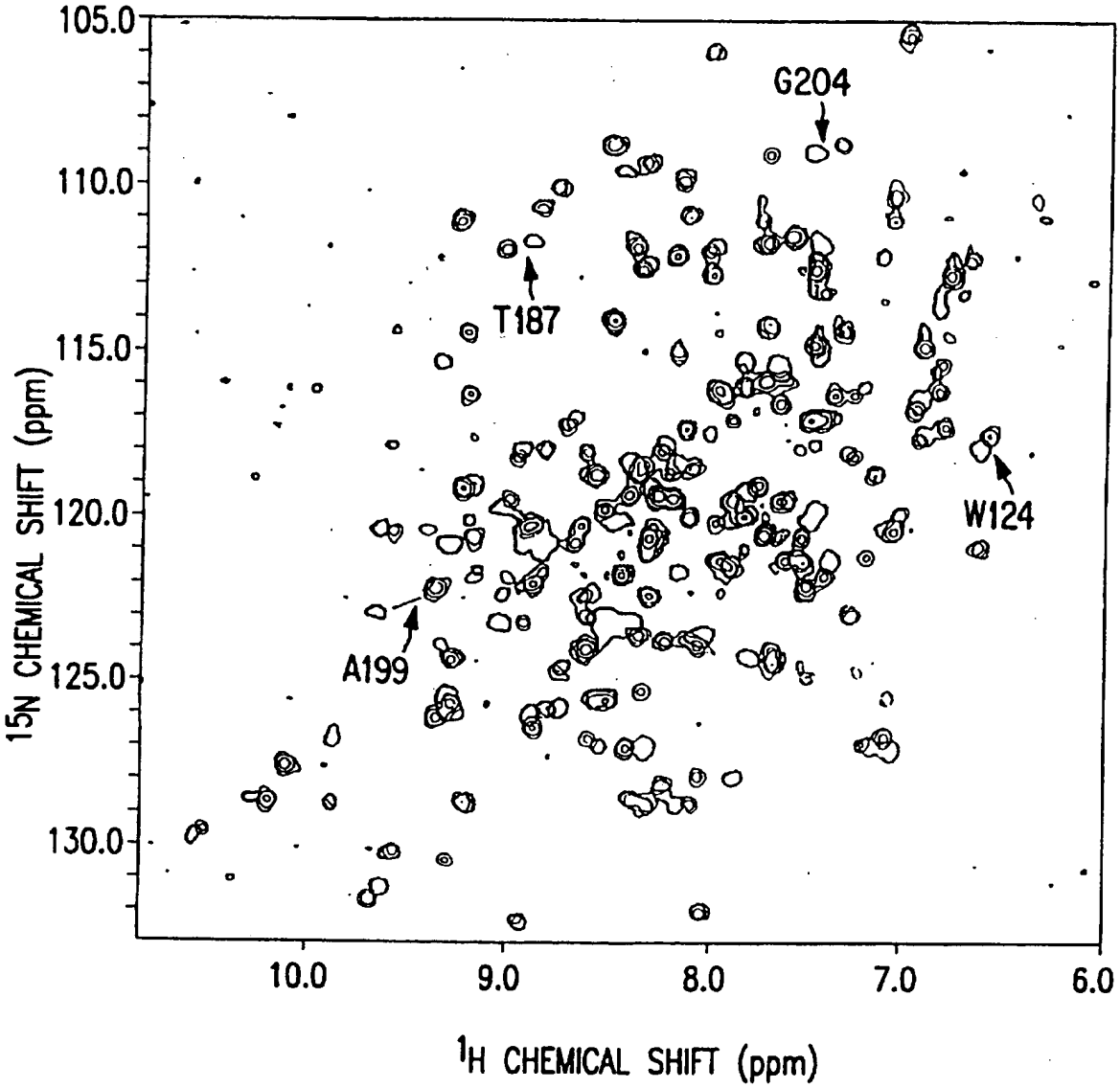

105.0
110.0
115.0
120.0
125.0
130.0
15N CHEMICAL SHIFT (ppm)
G204
T187
W124
A199
10.0
9.0
8.0
7.0
6.0
1H CHEMICAL SHIFT (ppm)

FIG.4

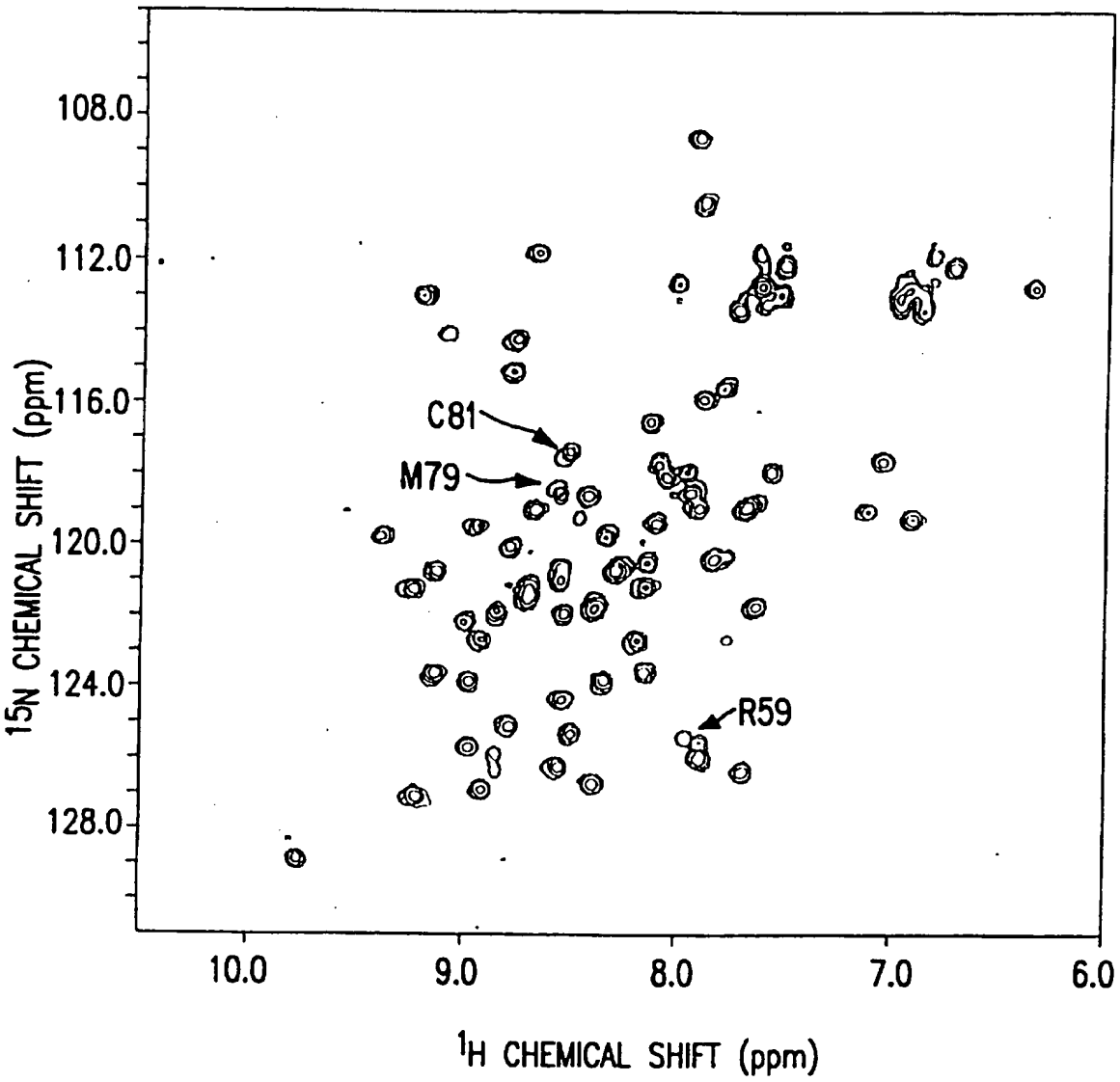

108.0
112.0
116.0
120.0
124.0
128.0
15N CHEMICAL SHIFT (ppm)
C81
M79
R59
10.0
9.0
8.0
7.0
6.0
1H CHEMICAL SHIFT (ppm)

FIG.5

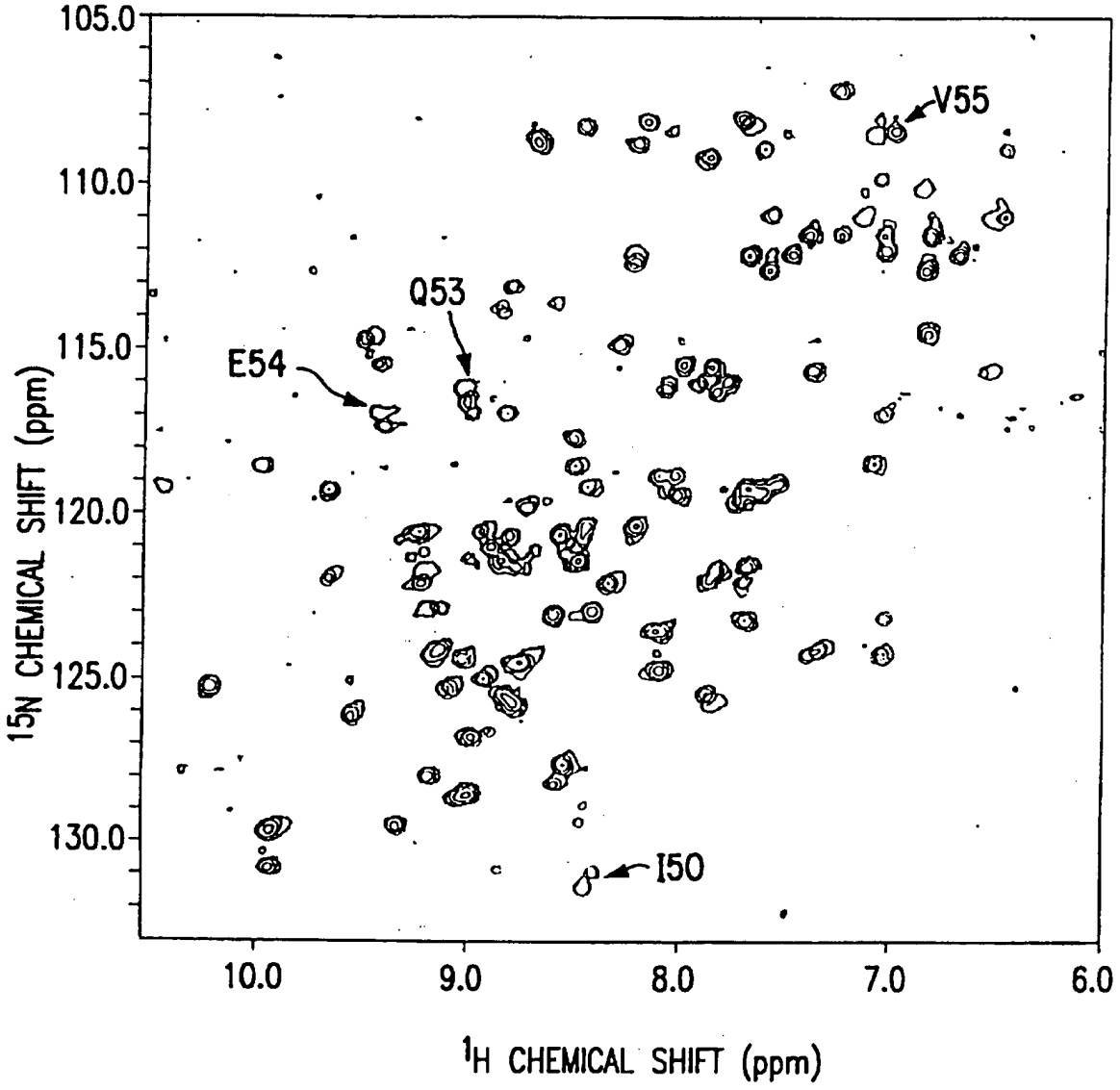

105.0
110.0
115.0
120.0
125.0
130.0
15N CHEMICAL SHIFT (ppm)
V55
Q53
E54
I50
10.0
9.0
8.0
7.0
6.0
1H CHEMICAL SHIFT (ppm)

FIG.6

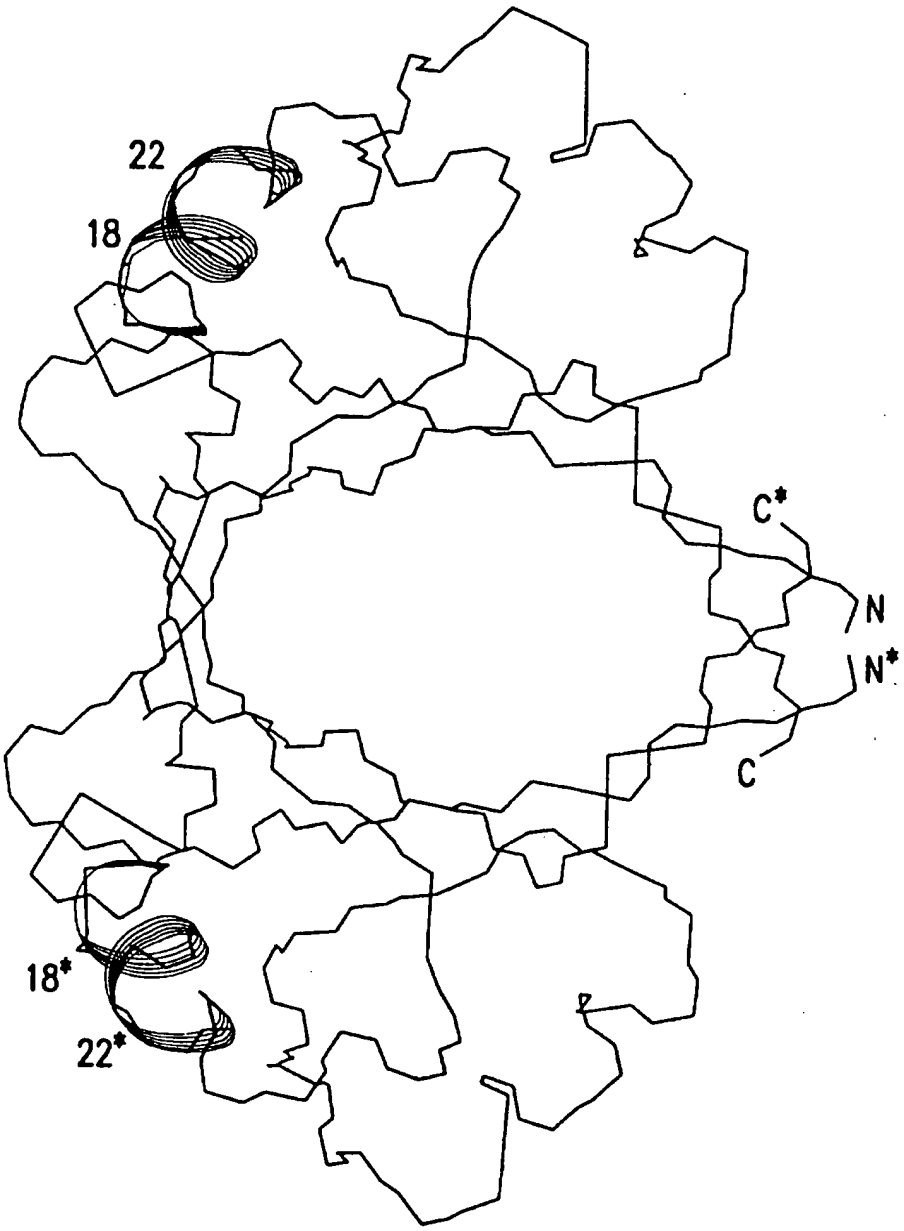


FIG.7

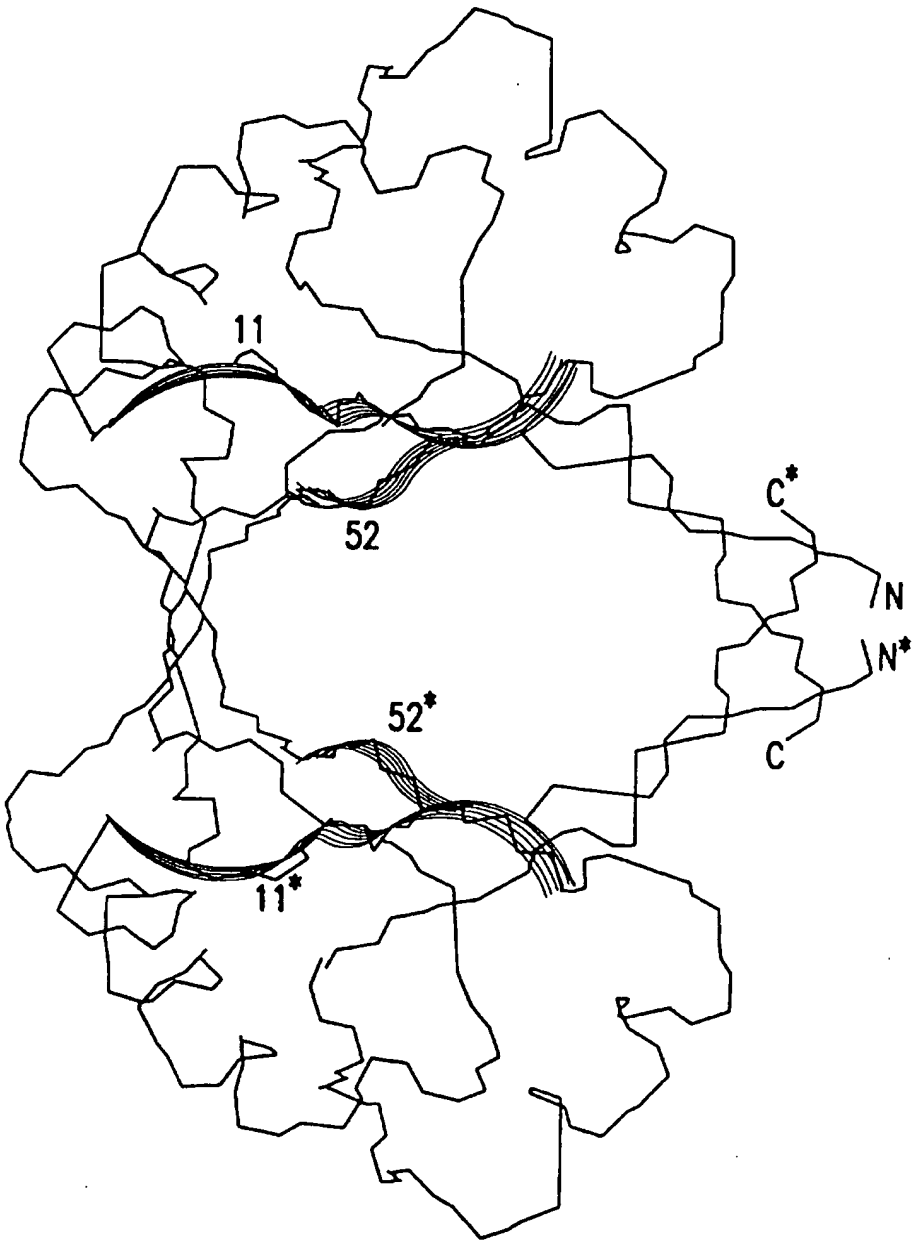


FIG.8

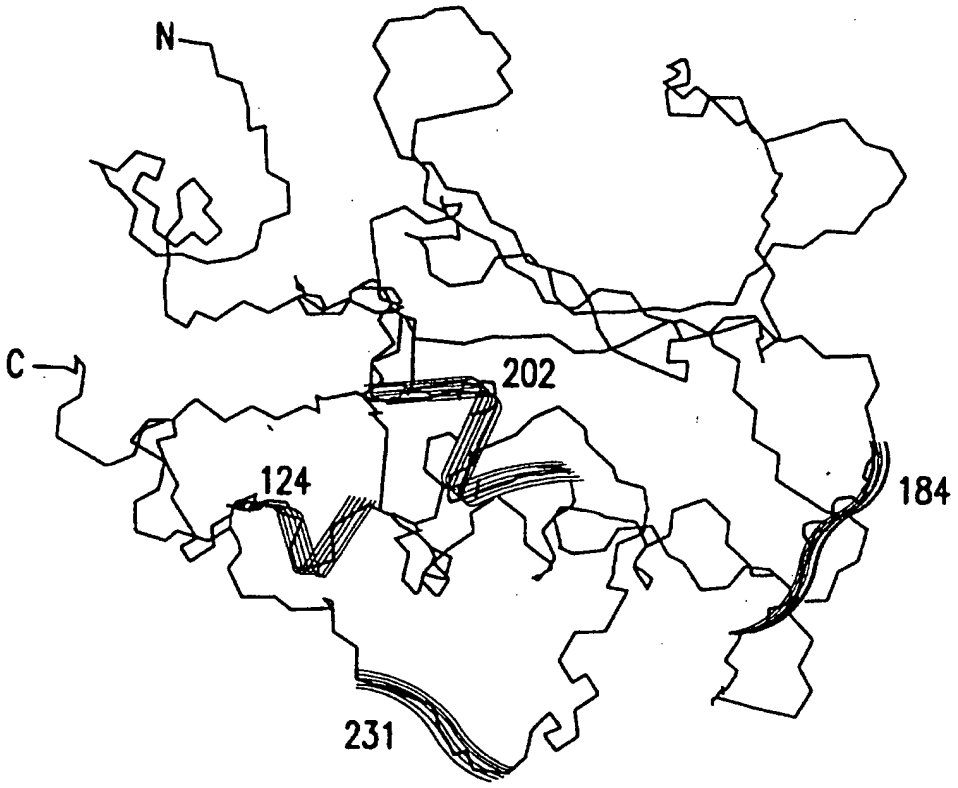


FIG.9

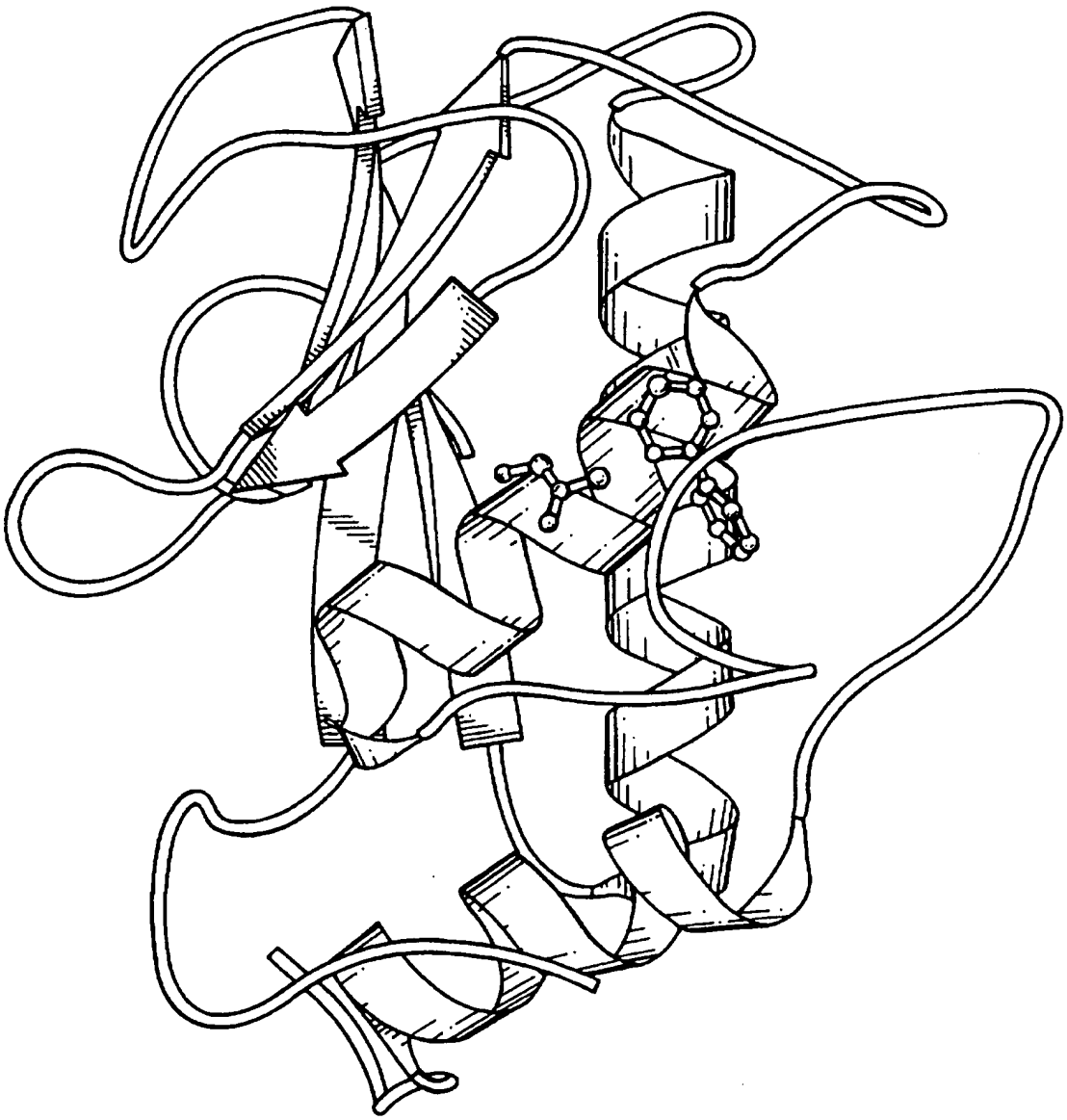


FIG. 10